(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 September 2001 (07.09.2001)

PCT

(10) International Publication Number
WO 01/64920 A2

(51) International Patent Classification[7]: C12N 15/62, C07K 14/22, 19/00

(21) International Application Number: PCT/IB01/00420

(22) International Filing Date: 28 February 2001 (28.02.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0004695.3 28 February 2000 (28.02.2000) GB
0027675.8 13 November 2000 (13.11.2000) GB

(71) Applicant *(for all designated States except US)*: **CHIRON SPA** [IT/IT]; Via Fiorentina, 1, I-53100 Siena (IT).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **ARICO, Maria, Beatrice** [IT/IT]; Chiron SpA, Via Fiorentina 1, I-53100 Siena (IT). **COMANDUCCI, Maurizio** [IT/IT]; Chiron SpA, Via Fiorentina 1, I-53100 Siena (IT). **GALEOTTI, Cesira** [IT/IT]; Chiron SpA, Via Fiorentina, 1, I-53100 Siena (IT). **MASIGNANI, Vega** [IT/IT]; Chiron SpA, Via Fiorentina, 1, I-53100 Siena (IT). **GIULIANI, Marzia, Monica** [IT/IT]; Chiron SpA, Via Fiorentina, 1, I-53100 Siena (IT). **PIZZA, Mariagrazia** [IT/IT]; Chiron SpA, Via Fiorentina, 1, I-53100 Siena (IT).

(74) Agents: **HALLYBONE, Huw, George** et al.; Carpmaels & Ransford, 43 Bloomsbury Square, London WC1A 2RA (GB).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

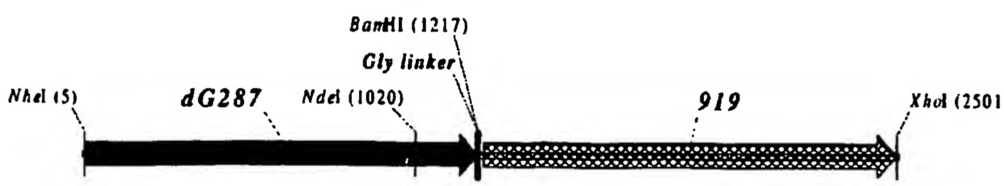
(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/64920 A2

(54) Title: HYBRID EXPRESSION OF NEISSERIAL PROTEINS



(57) Abstract: Two or more Neisserial proteins (e.g. A and B) are expressed as a single hybrid protein which can be represented simply by the formula NH_2-A-B-COOH.

HYBRID EXPRESSION OF NEISSERIAL PROTEINS

All documents cited herein are incorporated by reference in their entirety.

TECHNICAL FIELD

This invention is in the field of protein expression. In particular, it relates to the heterologous expression of proteins from *Neisseria* (*e.g. N.gonorrhoeae* or, preferably, *N.meningitidis*).

BACKGROUND ART

International patent applications WO99/24578, WO99/36544, WO99/57280 and WO00/22430 disclose proteins from *Neisseria meningitidis* and *Neisseria gonorrhoeae*. These proteins are typically described as being expressed in *E.coli* (*i.e.* heterologous expression) as either N-terminal GST-fusions or C-terminal His-tag fusions, although other expression systems, including expression in native *Neisseria*, are also disclosed.

It is an object of the present invention to provide alternative and improved approaches for the heterologous expression of these proteins. These approaches will typically affect the level of expression, the ease of purification, the cellular localisation of expression, and/or the immunological properties of the expressed protein.

DISCLOSURE OF THE INVENTION

In accordance with the invention, two or more (*e.g.* 3, 4, 5, 6 or more) proteins of the invention are expressed as a single hybrid protein. It is preferred that no non-Neisserial fusion partner (*e.g.* GST or poly-His) is used.

This offers two advantages. Firstly, a protein that may be unstable or poorly expressed on its own can be assisted by adding a suitable hybrid partner that overcomes the problem. Secondly, commercial manufacture is simplified – only one expression and purification need be employed in order to produce two separately-useful proteins.

Thus the invention provides a method for the simultaneous heterologous expression of two or more proteins of the invention, in which said two or more proteins of the invention are fused (*i.e.* they are translated as a single polypeptide chain).

The method will typically involve the steps of: obtaining a first nucleic acid encoding a first protein of the invention; obtaining a second nucleic acid encoding a second protein of the

invention; ligating the first and second nucleic acids. The resulting nucleic acid may be inserted into an expression vector, or may already be part of an expression vector.

Where just two proteins are joined, the hybrid protein can be represented simply by the formula NH_2-A—B-COOH. A and B can each be selected from any Neisserial proteins, and in particular those represented by SEQ#s 1-4326. The method is well suited to the expression of proteins orf1, orf4, orf25, orf40, Orf46/46.1, orf83, 233, 287, 292L, 564, 687, 741, 907, 919, 953, 961 and 983.

The 42 hybrids indicated by 'X' in the following table of form NH_2-A—B-COOH are preferred:

↓A B→	ORF46.1	287	741	919	953	961	983
ORF46.1		X	X	X	X	X	X
287	X		X	X	X	X	X
741	X	X		X	X	X	X
919	X	X	X		X	X	X
953	X	X	X	X		X	X
961	X	X	X	X	X		X
983	X	X	X	X	X	X	

Preferred proteins to be expressed as hybrids are thus ORF46.1, 287, 741, 919, 953, 961 and 983. These may be used in their essentially full-length form, or poly-glycine deletions (ΔG) forms may be used (*e.g.* ΔG-287, ΔGTbp2, ΔG741, ΔG983 *etc.*), or truncated forms may be used (*e.g.* Δ1-287, Δ2-287 *etc.*), or domain-deleted versions may be used (*e.g.* 287B, 287C, 287BC, $ORF46_{1-433}$, $ORF46_{433-608}$, ORF46, 961c *etc.*) and so on.

Particularly preferred are: (a) a hybrid protein comprising 919 and 287; (b) a hybrid protein comprising 953 and 287; (c) a hybrid protein comprising 287 and ORF46.1; (d) a hybrid protein comprising ORF1 and ORF46.1; (e) a hybrid protein comprising 919 and ORF46.1; (f) a hybrid protein comprising ORF46.1 and 919; (g) a hybrid protein comprising ORF46.1, 287 and 919; (h) a hybrid protein comprising 919 and 519; and (i) a hybrid protein comprising ORF97 and 225.

Further embodiments are shown in the drawings and include ΔG287-919, ΔG287-953, ΔG287-961, ΔG983-ORF46.1, ΔG983-741, ΔG983-961, ΔG983-961C, ΔG741-961, ΔG741-961C, ΔG741-983, ΔG741-ORF46.1, ORF46.1-741, ORF46.1-961, ORF46.1-961C,

961-ORF46.1, 961-741, 961-983, 961C-ORF46.1, 961C-741, 961C-983, 961CL–ORF46.1, 961CL-741, and 961CL-983.

Where 287 is used, it is preferably at the C-terminal end of a hybrid; if it is to be used at the N-terminus, if is preferred to use a ΔG form of 287 is used (*e.g.* as the N-terminus of a hybrid with ORF46.1, 919, 953 or 961).

Where 287 is used, this is preferably from strain 2996 or from strain 394/98.

Where 961 is used, this is preferably at the N-terminus. Domain forms of 961 may be used.

Alignments of polymorphic forms of ORF46, 287, 919 and 953 are disclosed in WO00/66741. Any of these polymorphs can be used according to the present invention.

Preferably, the constituent proteins (A and B) in a hybrid protein according to the invention will be from the same strain.

The fused proteins in the hybrid may be joined directly, or may be joined via a linker peptide *e.g.* via a poly-glycine linker (*i.e.* G_n where $n = 3, 4, 5, 6, 7, 8, 9, 10$ or more) or via a short peptide sequence which facilitates cloning. It is evidently preferred not to join a ΔG protein to the C-terminus of a poly-glycine linker.

The fused proteins may lack native leader peptides or may include the leader peptide sequence of the N-terminal fusion partner.

Host

It is preferred to utilise a heterologous host. The heterologous host may be prokaryotic or eukaryotic. It is preferably *E.coli*, but other suitable hosts include *Bacillus subtilis, Vibrio cholerae, Salmonella typhi, Salmonenna typhimurium, Neisseria meningitidis, Neisseria gonorrhoeae, Neisseria lactamica, Neisseria cinerea, Mycobateria* (*e.g. M.tuberculosis*), yeast *etc.*

***Vectors, hosts* etc.**

As well as the methods described above, the invention provides (a) nucleic acid and vectors useful in these methods (b) host cells containing said vectors (c) proteins expressed or expressable by the methods (d) compositions comprising these proteins, which may be suitable as vaccines, for instance, or as diagnostic reagents, or as immunogenic compositions (e) these compositions for use as medicaments (*e.g.* as vaccines) or as diagnostic reagents (f)

the use of these compositions in the manufacture of (1) a medicament for treating or preventing infection due to Neisserial bacteria (2) a diagnostic reagent for detecting the presence of Neisserial bacteria or of antibodies raised against Neisserial bacteria, and/or (3) a reagent which can raise antibodies against Neisserial bacteria and (g) a method of treating a patient, comprising administering to the patient a therapeutically effective amount of these compositions.

Sequences

The invention also provides a protein or a nucleic acid having any of the sequences set out in the following examples. It also provides proteins and nucleic acid having sequence identity to these. As described above, the degree of 'sequence identity' is preferably greater than 50% (eg. 60%, 70%, 80%, 90%, 95%, 99% or more).

Nomenclature herein

The 2166 protein sequences disclosed in WO99/24578, WO99/36544 and WO99/57280 are referred to herein by the following SEQ# numbers:

Application	Protein sequences	SEQ# herein
WO99/24578	Even SEQ IDs 2-892	SEQ#s 1-446
WO99/36544	Even SEQ IDs 2-90	SEQ#s 447-491
WO99/57280	Even SEQ IDs 2-3020 Even SEQ IDs 3040-3114 SEQ IDs 3115-3241	SEQ#s 492-2001 SEQ#s 2002-2039 SEQ#s 2040-2166

In addition to this SEQ# numbering, the naming conventions used in WO99/24578, WO99/36544 and WO99/57280 are also used (*e.g.* 'ORF4', 'ORF40', 'ORF40-1' *etc.* as used in WO99/24578 and WO99/36544; 'm919', 'g919' and 'a919' *etc.* as used in WO99/57280).

The 2160 proteins NMB0001 to NMB2160 from Tettelin *et al.* [*Science* (2000) 287:1809-1815] are referred to herein as SEQ#s 2167-4326 [see also WO00/66791].

The term 'protein of the invention' as used herein refers to a protein comprising:

(a) one of sequences SEQ#s 1-4326; or

(b) a sequence having sequence identity to one of SEQ#s 1-4326; or

(c) a fragment of one of SEQ#s 1-4326.

The degree of 'sequence identity' referred to in (b) is preferably greater than 50% (*eg.* 60%, 70%, 80%, 90%, 95%, 99% or more). This includes mutants and allelic variants [*e.g.* see WO00/66741]. Identity is preferably determined by the Smith-Waterman homology search algorithm as implemented in the MPSRCH program (Oxford Molecular), using an affine gap search with parameters *gap open penalty=12* and *gap extension penalty=1*. Typically, 50% identity or more between two proteins is considered to be an indication of functional equivalence.

The 'fragment' referred to in (c) should comprise at least *n* consecutive amino acids from one of SEQ#s 1-4326 and, depending on the particular sequence, *n* is 7 or more (*eg.* 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100 or more). Preferably the fragment comprises an epitope from one of SEQ#s 1-4326. Preferred fragments are those disclosed in WO00/71574 and WO01/04316.

Preferred proteins of the invention are found in *N.meningitidis* serogroup B.

Preferred proteins for use according to the invention are those of serogroup B *N.meningitidis* strain 2996 or strain 394/98 (a New Zealand strain). Unless otherwise stated, proteins mentioned herein are from *N.meningitidis* strain 2996. It will be appreciated, however, that the invention is not in general limited by strain. References to a particular protein (*e.g.* '287', '919' *etc.*) may be taken to include that protein from any strain.

It will be appreciated that references to "nucleic acid" includes DNA and RNA, and also their analogues, such as those containing modified backbones, and also peptide nucleic acids (PNA) *etc.*

BRIEF DESCRIPTION OF DRAWINGS

Figures 1 to 26 show hybrid proteins according to the invention.

MODES FOR CARRYING OUT THE INVENTION

Example 1 – hybrids of ORF46

The complete ORF46 protein from *N.meningitidis* (serogroup B, strain 2996) has the following sequence:

```
  1 LGISRKISLI LSILAVCLPM HAHASDLAND SFIRQVLDRQ HFEPDGKYHL
 51 FGSRGELAER SGHIGLGKIQ SHQLGNLMIQ QAAIKGNIGY IVRFSDHGHE
101 VHSPFDNHAS HSDSDEAGSP VDGFSLYRIH WDGYEHHPAD GYDGPQGGGY
```

```
151  PAPKGARDIY SYDIKGVAQN IRLNLTDNRS TGQRLADRFH NAGSMLTQGV
201  GDGFKRATRY SPELDRSGNA AEAFNGTADI VKNIIGAAGE IVGAGDAVQG
251  ISEGSNIAVM HGLGLLSTEN KMARINDLAD MAQLKDYAAA AIRDWAVQNP
301  NAAQGIEAVS NIFMAAIPIK GIGAVRGKYG LGGITAHPIK RSQMGAIALP
351  KGKSAVSDNF ADAAYAKYPS PYHSRNIRSN LEQRYGKENI TSSTVPPSNG
401  KNVKLADQRH PKTGVPFDGK GFPNFEKHVK YDTKLDIQEL SGGGIPKAKP
451  VSDAKPRWEV DRKLNKLTTR EQVEKNVQEI RNGNKNSNFS QHAQLEREIN
501  KLKSADEINF ADGMGKFTDS MNDKAFSRLV KSVKENGFTN PVVEYVEING
551  KAYIVRGNNR VFAAEYLGRI HELKFKKVDF PVPNTSWKNP TDVLNESGNV
601  KRPRYRSK*
```

The leader peptide is underlined.

The sequences of ORF46 from other strains can be found in WO00/66741.

ORF46 has been fused at its C-terminus and N-terminus with 287, 919, and ORF1. The hybrid proteins were generally insoluble, but gave some good ELISA and bactericidal results (against the homologous 2996 strain):

Protein	ELISA	Bactericidal Ab
Orf1-Orf46.1-His	850	256
919-Orf46.1-His	12900	512
919-287-Orf46-His	n.d.	n.d.
Orf46.1-287His	150	8192
Orf46.1-919His	2800	2048
Orf46.1-287-919His	3200	16384

For comparison, 'triple' hybrids of ORF46.1, 287 (either as a GST fusion, or in ΔG287 form) and 919 were constructed and tested against various strains (including the homologous 2996 strain) *versus* a simple mixture of the three antigens. FCA was used as adjuvant:

	2996	BZ232	MC58	NGH38	F6124	BZ133
Mixture	8192	256	512	1024	>2048	>2048
ORF46.1-287-919his	16384	256	4096	8192	8192	8192
ΔG287-919-ORF46.1his	8192	64	4096	8192	8192	16384
ΔG287-ORF46.1-919his	4096	128	256	8192	512	1024

Again, the hybrids show equivalent or superior immunological activity.

Hybrids of two proteins (strain 2996) were compared to the individual proteins against various heterologous strains:

	1000	MC58	F6124 (MenA)
ORF46.1-His	<4	4096	<4
ORF1-His	8	256	128
ORF1—ORF46.1-His	1024	512	1024

Again, the hybrid shows equivalent or superior immunological activity.

Example 2 – hybrids of ΔG287

The deletion of the $(Gly)_6$ sequence in 287 was found to have a dramatic effect on protein expression. The protein lacking the N-terminal amino acids up to GGGGGG is called 'ΔG287'. In strain MC58, its basic sequence (leader peptide underlined) is:

```
      SPDVKS ADTLSKPAAP VVSEKETEAK EDAPQAGSQG QGAPSAQGSQ DMAAVSEENT
GNGGAVTADN PKNEDEVAQN DMPQNAAGTD SSTPNHTPDP NMLAGNMENQ ATDAGESSQP
ANQPDMANAA DGMQGDDPSA GGQNAGNTAA QGANQAGNNQ AAGSSDPIPA SNPAPANGGS
NFGRVDLANG VLIDGPSQNI TLTHCKGDSC SGNNFLDEEV QLKSEFEKLS DADKISNYKK
DGKNDKFVGL VADSVQMKGI NQYIIFYKPK PTSFARFRRS ARSRRSLPAE MPLIPVNQAD
TLIVDGEAVS LTGHSGNIFA PEGNYRYLTY GAEKLPGGSY ALRVQGEPAK GEMLAGAAVY
NGEVLHFHTE NGRPYPTRGR FAAKVDFGSK SVDGIIDSGD DLHMGTQKFK AAIDGNGFKG
TWTENGSGDV SGKFYGPAGE EVAGKYSYRP TDAEKGGFGV FAGKKEQD*
```

ΔG287, with or without His-tag ('ΔG287-His' and 'ΔG287K', respectively), are expressed at very good levels in comparison with the '287-His' or '287 ^{untagged}'.

On the basis of gene variability data, variants of ΔG287-His were expressed in *E.coli* from a number of MenB strains, in particular from strains 2996, MC58, 1000, and BZ232. The results were also good – each of these gave high ELISA titres and also serum bactericidal titres of >8192. ΔG287K, expressed from pET-24b, gave excellent titres in ELISA and the serum bactericidal assay.

Deletion of poly-Gly sequences is also applicable to Tbp2 (NMB0460), 741 (NMB 1870) and 983 (NMB1969). When cloned in pET vector and expressed in *E.coli* without the sequence coding for their leader peptides and without poly-Gly (*i.e.* as "ΔG forms"), the same effect was seen – expression was good in the clones carrying the deletion of the poly-glycine stretch, and poor or absent if the glycines were present in the expressed protein.

ΔG287 was fused directly in-frame upstream of 919, 953, 961 (sequences shown below) and ORF46.1:

ΔG287–919

```
ATGGCTAGCCCCGATGTTAAATCGGCGGACACGCTGTCAAAACCGGCCGCTCCTGTTGTTGCTGAAAAAGAGACAGAG
GTAAAAGAAGATGCGCCACAGGCAGGTTCTCAAGGACAGGGCGCGCCATCCACACAAGGCAGCCAAGATATGGCGGCA
GTTTCGGCAGAAAATACAGGCAATGGCGGTGCGGCAACAACGGACAAACCCAAAAATGAAGACGAGGGACCGCAAAAT
GATATGCCGCAAAATTCCGCCGAATCCGCAAATCAAACAGGGAACAACCAACCCGCCGATTCTTCAGATTCCGCCCCC
GCGTCAAACCCTGCACCTGCGAATGGCGGTAGCAATTTTGGAAGGGTTGATTTGGCTAATGGCGTTTTGATTGATGGG
CCGTCGCAAAATATAACGTTGACCCACTGTAAAGGCGATTCTTGTAATGGTGATAATTTATTGGATGAAGAAGCACCG
TCAAAATCAGAATTTGAAAATTTAAATGAGTCTGAACGAATTGAGAAATATAAGAAAGATGGGAAAAGCGATAAATTT
ACTAATTTGGTTGCGACAGCAGTTCAAGCTAATGGAACTAACAAATATGTCATCATTTATAAAGACAAGTCCGCTTCA
TCTTCATCTGCGCGATTCAGGCGTTCTGCACGGTCGAGGAGGTCGCTTCCTGCCGAGATGCCGCTAATCCCCGTCAAT
CAGGCGGATACGCTGATTGTCGATGGGGAAGCGGTCAGCCTGACGGGGCATTCCGGCAATATCTTCGCGCCCGAAGGG
AATTACCGGTATCTGACTTACGGGGCGGAAAAATTGCCCGGCGGATCGTATGCCCTCCGTGTGCAAGGCGAACCGGCA
AAAGGCGAAATGCTTGCTGGCACGGCCGTGTACAACGGCGAAGTGCTGCATTTTCATACGGAAAACGGCCGTCCGTAC
CCGACTAGAGGCAGGTTTGCCGCAAAAGTCGATTTCGGCAGCAAATCTGTGGACGGCATTATCGACAGCGGCGATGAT
TTGCATATGGGTACGCAAAAATTCAAAGCCGCCATCGATGGAAACGGCTTTAAGGGGACTTGGACGGAAAATGGCGGC
GGGGATGTTTCCGGAAGGTTTTACGGCCCGGCCGGCGAGGAAGTGGCGGGAAAATACAGCTATCGCCCGACAGATGCG
```

GAAAAGGGCGGATTCGGCGTGTTTGCCGGCAAAAAAGAGCAGGATGGATCCGGAGGAGGAGGATGCCAAAGCAAGAGC
ATCCAAACCTTTCCGCAACCCGACACATCCGTCATCAACGGCCCGGACCGGCCGGTCGGCATCCCCGACCCCGCCGGA
ACGACGGTCGGCGGCGGCGGGGCCGTCTATACCGTTGTACCGCACCTGTCCCTGCCCCACTGGGCGGCGCAGGATTTC
GCCAAAAGCCTGCAATCCTTCCGCCTCGGCTGCGCCAATTTGAAAAACCGCCAAGGCTGGCAGGATGTGTGCGCCCAA
GCCTTTCAAACCCCCGTCCATTCCTTTCAGGCAAAACAGTTTTTTGAACGCTATTTCACGCCGTGGCAGGTTGCAGGC
AACGGAAGCCTTGCCGGTACGGTTACCGGCTATTACGAGCCGGTGCTGAAGGGCGACGACAGGCGGACGGCACAAGCC
CGCTTCCCGATTTACGGTATTCCCGACGATTTTATCTCCGTCCCCCTGCCTGCCGGTTTGCGGAGCGGAAAAGCCCTT
GTCCGCATCAGGCAGACGGGAAAAAACAGCGGCACAATCGACAATACCGGCGGCACACATACCGCCGACCTCTCCCGA
TTCCCCATCACCGCGCGCACAACGGCAATCAAAGGCAGGTTTGAAGGAAGCCGCTTCCTCCCCTACCACACGCGCAAC
CAAATCAACGGCGGCGCGCTTGACGGCAAAGCCCCGATACTCGGTTACGCCGAAGACCCCGTCGAACTTTTTTTTATG
CACATCCAAGGCTCGGGCCGTCTGAAAACCCCGTCCGGCAAATACATCCGCATCGGCTATGCCGACAAAAACGAACAT
CCCTACGTTTCCATCGGACGCTATATGGCGGACAAAGGCTACCTCAAGCTCGGGCAGACCTCGATGCAGGGCATCAAA
GCCTATATGCGGCAAAATCCGCAACGCCTCGCCGAAGTTTTGGGTCAAAACCCCAGCTATATCTTTTTCCGCGAGCTT
GCCGGAAGCAGCAATGACGGTCCCGTCGGCGCACTGGGCACGCCGTTGATGGGGGAATATGCCGGCGCAGTCGACCGG
CACTACATTACCTTGGGCGCGCCCTTATTTGTCGCCACCGCCCATCCGGTTACCCGCAAAGCCCTCAACCGCCTGATT
ATGGCGCAGGATACCGGCAGCGCGATTAAAGGCGCGGTGCGCGTGGATTATTTTTGGGGATACGGCGACGAAGCCGGC
GAACTTGCCGGCAAACAGAAAACCACGGGTTACGTCTGGCAGCTCCTACCCAACGGTATGAAGCCCGAATACCGCCCG
TAACTCGAG

```
  1   MASPDVKSAD TLSKPAAPVV AEKETEVKED APQAGSQGQG APSTQGSQDM
 51   AAVSAENTGN GGAATTDKPK NEDEGPQNDM PQNSAESANQ TGNNQPADSS
101   DSAPASNPAP ANGGSNFGRV DLANGVLIDG PSQNITLTHC KGDSCNGDNL
151   LDEEAPSKSE FENLNESERI EKYKKDGKSD KFTNLVATAV QANGTNKYVI
201   IYKDKSASSS SARFRRSARS RRSLPAEMPL IPVNQADTLI VDGEAVSLTG
251   HSGNIFAPEG NYRYLTYGAE KLPGGSYALR VQGEPAKGEM LAGTAVYNGE
301   VLHFHTENGR PYPTRGRFAA KVDFGSKSVD GIIDSGDDLH MGTQKFKAAI
351   DGNGFKGTWT ENGGGDVSGR FYGPAGEEVA GKYSYRPTDA EKGGFGVFAG
401   KKEQDGSGGG GCQSKSIQTF PQPDTSVING PDRPVGIPDP AGTTVGGGGA
451   VYTVVPHLSL PHWAAQDFAK SLQSFRLGCA NLKNRQGWQD VCAQAFQTPV
501   HSFQAKQFFE RYFTPWQVAG NGSLAGTVTG YYEPVLKGDD RRTAQARFPI
551   YGIPDDFISV PLPAGLRSGK ALVRIRQTGK NSGTIDNTGG THTADLSRFP
601   ITARTTAIKG RFEGSRFLPY HTRNQINGGA LDGKAPILGY AEDPVELFFM
651   HIQGSGRLKT PSGKYIRIGY ADKNEHPYVS IGRYMADKGY LKLGQTSMQG
701   IKAYMRQNPQ RLAEVLGQNP SYIFFRELAG SSNDGPVGAL GTPLMGEYAG
751   AVDRHYITLG APLFVATAHP VTRKALNRLI MAQDTGSAIK GAVRVDYFWG
801   YGDEAGELAG KQKTTGYVWQ LLPNGMKPEY RP*
```

ΔG287–953

ATGGCTAGCCCCGATGTTAAATCGGCGGACACGCTGTCAAAACCGGCCGCTCCTGTTGTTGCTGAAAAAGAGACAGAG
GTAAAAGAAGATGCGCCACAGGCAGGTTCTCAAGGACAGGGCGCGCCATCCACACAAGGCAGCCAAGATATGGCGGCA
GTTTCGGCAGAAAATACAGGCAATGGCGGTGCGGCAACAACGGACAAACCCAAAAATGAAGACGAGGGACCGCAAAAT
GATATGCCGCAAAATTCCGCCGAATCCGCAAATCAAACAGGGAACAACCAACCCGCCGATTCTTCAGATTCCGCCCCC
GCGTCAAACCCTGCACCTGCGAATGGCGGTAGCAATTTTGGAAGGGTTGATTTGGCTAATGGCGTTTTGATTGATGGG
CCGTCGCAAAATATAACGTTGACCCACTGTAAAGGCGATTCTTGTAATGGTGATAATTTATTGGATGAAGAAGCACCG
TCAAAATCAGAATTTGAAAATTTAAATGAGTCTGAACGAATTGAGAAATATAAGAAAGATGGGAAAAGCGATAAATTT
ACTAATTTGGTTGCGACAGCAGTTCAAGCTAATGGAACTAACAAATATGTCATCATTTATAAAGACAAGTCCGCTTCA
TCTTCATCTGCGCGATTCAGGCGTTCTGCACGGTCGAGGAGGTCGCTTCCTGCCGAGATGCCGCTAATCCCCGTCAAT
CAGGCGGATACGCTGATTGTCGATGGGGAAGCGGTCAGCCTGACGGGGCATTCCGGCAATATCTTCGCGCCCGAAGGG
AATTACCGGTATCTGACTTACGGGGCGGAAAAATTGCCCGGCGGATCGTATGCCCTCCGTGTGCAAGGCGAACCGGCA
AAAGGCGAAATGCTTGCTGGCACGGCCGTGTACAACGGCGAAGTGCTGCATTTTCATACGGAAAACGGCCGTCCGTAC
CCGACTAGAGGCAGGTTTGCCGCAAAAGTCGATTTCGGCAGCAAATCTGTGGACGGCATTATCGACAGCGGCGATGAT
TTGCATATGGGTACGCAAAAATTCAAAGCCGCCATCGATGGAAACGGCTTTAAGGGGACTTGGACGGAAAATGGCGGC
GGGGATGTTTCCGGAAGGTTTTACGGCCCGGCCGGCGAGGAAGTGGCGGGAAAATACAGCTATCGCCCGACAGATGCG
GAAAAGGGCGGATTCGGCGTGTTTGCCGGCAAAAAAGAGCAGGATGGATCCGGAGGAGGAGCCACCTACAAAGTG
GACGAATATCACGCCAACGCCCGTTTCGCCATCGACCATTTCAACACCAGCACCAACGTCGGCGGTTTTTACGGTCTG
ACCGGTTCCGTCGAGTTCGACCAAGCAAAACGCGACGGTAAAATCGACATCACCATCCCCGTTGCCAACCTGCAAAGC
GGTTCGCAACACTTTACCGACCACCTGAAATCAGCCGACATCTTCGATGCCGCCCAATATCCGGACATCCGCTTTGTT
TCCACCAAATTCAACTTCAACGGCAAAAAACTGGTTTCCGTTGACGGCAACCTGACCATGCACGGCAAAACCGCCCCC
GTCAAACTCAAAGCCGAAAAATTCAACTGCTACCAAAGCCCGATGGCGAAAACCGAAGTTTGCGGCGGCGACTTCAGC
ACCACCATCGACCGCACCAAATGGGGCGTGGACTACCTCGTTAACGTTGGTATGACCAAAAGCGTCCGCATCGACATC
CAAATCGAGGCAGCCAAACAATAACTCGAG

```
  1   MASPDVKSAD TLSKPAAPVV AEKETEVKED APQAGSQGQG APSTQGSQDM
 51   AAVSAENTGN GGAATTDKPK NEDEGPQNDM PQNSAESANQ TGNNQPADSS
101   DSAPASNPAP ANGGSNFGRV DLANGVLIDG PSQNITLTHC KGDSCNGDNL
151   LDEEAPSKSE FENLNESERI EKYKKDGKSD KFTNLVATAV QANGTNKYVI
```

201 IYKDKSASSS SARFRRSARS RRSLPAEMPL IPVNQADTLI VDGEAVSLTG
251 HSGNIFAPEG NYRYLTYGAE KLPGGSYALR VQGEPAKGEM LAGTAVYNGE
301 VLHFHTENGR PYPTRGRFAA KVDFGSKSVD GIIDSGDDLH MGTQKFKAAI
351 DGNGFKGTWT ENGGGDVSGR FYGPAGEEVA GKYSYRPTDA EKGGFGVFAG
401 KKEQDGSGGG GATYKVDEYH ANARFAIDHF NTSTNVGGFY GLTGSVEFDQ
451 AKRDGKIDIT IPVANLQSGS QHFTDHLKSA DIFDAAQYPD IRFVSTKFNF
501 NGKKLVSVDG NLTMHGKTAP VKLKAEKFNC YQSPMAKTEV CGGDFSTTID
551 RTKWGVDYLV NVGMTKSVRI DIQIEAAKQ*

ΔG287-961

ATGGCTAGCCCCGATGTTAAATCGGCGGACACGCTGTCAAAACCGGCCGCTCCTGTTGTTGCTGAAAAAGAGACAGAG
GTAAAAGAAGATGCGCCACAGGCAGGTTCTCAAGGACAGGGCGCGCCATCCACACAAGGCAGCCAAGATATGGCGGCA
GTTTCGGCAGAAAATACAGGCAATGGCGGTGCGGCAACAACGGACAAACCCAAAAATGAAGACGAGGGACCGCAAAAT
GATATGCCGCAAAATTCCGCCGAATCCGCAAATCAAACAGGGAACAACCAACCCGCCGATTCTTCAGATTCCGCCCCC
GCGTCAAACCCTGCACCTGCGAATGGCGGTAGCAATTTTGGAAGGGTTGATTTGGCTAATGGCGTTTTGATTGATGGG
CCGTCGCAAAATATAACGTTGACCCACTGTAAAGGCGATTCTTGTAATGGTGATAATTTATTGGATGAAGAAGCACCG
TCAAAATCAGAATTTGAAAATTTAAATGAGTCTGAACGAATTGAGAAATATAAGAAAGATGGGAAAAGCGATAAATTT
ACTAATTTGGTTGCGACAGCAGTTCAAGCTAATGGAACTAACAAATATGTCATCATTTATAAAGACAAGTCCGCTTCA
TCTTCATCTGCGCGATTCAGGCGTTCTGCACGGTCGAGGAGGTCGCTTCCTGCCGAGATGCCGCTAATCCCCGTCAAT
CAGGCGGATACGCTGATTGTCGATGGGGAAGCGGTCAGCCTGACGGGGCATTCCGGCAATATCTTCGCGCCCGAAGGG
AATTACCGGTATCTGACTTACGGGGCGGAAAAATTGCCCGGCGGATCGTATGCCCTCCGTGTGCAAGGCGAACCGGCA
AAAGGCGAAATGCTTGCTGGCACGGCCGTGTACAACGGCGAAGTGCTGCATTTTCATACGGAAAACGGCCGTCCGTAC
CCGACTAGAGGCAGGTTTGCCGCAAAAGTCGATTTCGGCAGCAAATCTGTGGACGGCATTATCGACAGCGGCGATGAT
TTGCATATGGGTACGCAAAAATTCAAAGCCGCCATCGATGGAAACGGCTTTAAGGGGACTTGGACGGAAAATGGCGGC
GGGGATGTTTCCGGAAGGTTTTACGGCCCGGCCGGCGAGGAAGTGGCGGGAAAATACAGCTATCGCCCGACAGATGCG
GAAAAGGGCGGATTCGGCGTGTTTGCCGGCAAAAAAGAGCAGGATGGATCCGGAGGAGGAGGAGCCACAAACGACGAC
GATGTTAAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATCAACGGTTTCAAAGCTGGA
GAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCCGATGTTGAAGCCGACGAC
TTTAAAGGTCTGGGTCTGAAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAACAAACAAAACGTCGATGCC
AAAGTAAAAGCTGCAGAATCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGATGCCGCTTTAGCAGATACT
GATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAATTGGGAGAAAATATAACGACATTTGCTGAAGAGACTAAG
ACAAATATCGTAAAAATTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCATGCCGAAGCATTCAACGAT
ATCGCCGATTCATTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAATGAAGCCAAACAGACGGCC
GAAGAAACCAAACAAAACGTCGATGCCAAAGTAAAAGCTGCAGAAACTGCAGCAGGCAAAGCCGAAGCTGCCGCTGGC
ACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTCGCTGCAAAAGTTACCGACATCAAAGCTGATATCGCTACGAAC
AAAGATAATATTGCTAAAAAAGCAAACAGTGCCGACGTGTACACCAGAGAAGAGTCTGACAGCAAATTTGTCAGAATT
GATGGTCTGAACGCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCCATTGCCGATCACGATACT
CGCCTGAACGGTTTGGATAAAACAGTGTCAGACCTGCGCAAAGAAACCCGCCAAGGCCTTGCAGAACAAGCCGCGCTC
TCCGGTCTGTTCCAACCTTACAACGTGGGTCGGTTCAATGTAACGGCTGCAGTCGGCGGCTACAAATCCGAATCGGCA
GTCGCCATCGGTACCGGCTTCCGCTTTACCGAAAACTTTGCCGCCAAAGCAGGCGTGGCAGTCGGCACTTCGTCCGGT
TCTTCCGCAGCCTACCATGTCGGCGTCAATTACGAGTGGTAACTCGAG

1 MASPDVKSAD TLSKPAAPVV AEKETEVKED APQAGSQGQG APSTQGSQDM
51 AAVSAENTGN GGAATTDKPK NEDEGPQNDM PQNSAESANQ TGNNQPADSS
101 DSAPASNPAP ANGGSNFGRV DLANGVLIDG PSQNITLTHC KGDSCNGDNL
151 LDEEAPSKSE FENLNESERI EKYKKDGKSD KFTNLVATAV QANGTNKYVI
201 IYKDKSASSS SARFRRSARS RRSLPAEMPL IPVNQADTLI VDGEAVSLTG
251 HSGNIFAPEG NYRYLTYGAE KLPGGSYALR VQGEPAKGEM LAGTAVYNGE
301 VLHFHTENGR PYPTRGRFAA KVDFGSKSVD GIIDSGDDLH MGTQKFKAAI
351 DGNGFKGTWT ENGGGDVSGR FYGPAGEEVA GKYSYRPTDA EKGGFGVFAG
401 KKEQDGSGGG GATNDDDVKK AATVAIAAAY NNGQEINGFK AGETIYDIDE
451 DGTITKKDAT AADVEADDFK GLGLKKVVTN LTKTVNENKQ NVDAKVKAAE
501 SEIEKLTTKL ADTDAALADT DAALDATTNA LNKLGENITT FAEETKTNIV
551 KIDEKLEAVA DTVDKHAEAF NDIADSLDET NTKADEAVKT ANEAKQTAEE
601 TKQNVDAKVK AAETAAGKAE AAAGTANTAA DKAEAVAAKV TDIKADIATN
651 KDNIAKKANS ADVYTREESD SKFVRIDGLN ATTEKLDTRL ASAEKSIADH
701 DTRLNGLDKT VSDLRKETRQ GLAEQAALSG LFQPYNVGRF NVTAAVGGYK
751 SESAVAIGTG FRFTENFAAK AGVAVGTSSG SSAAYHVGVN YEW*

	ELISA	Bactericidal
ΔG287-953-His	3834	65536
ΔG287-961-His	108627	65536

The bactericidal efficacy (homologous strain) of antibodies raised against the hybrid proteins was compared with antibodies raised against simple mixtures of the component antigens (using 287-GST) for 919 and ORF46.1:

	Mixture with 287	Hybrid with ΔG287
919	32000	128000
ORF46.1	128	16000

Data for bactericidal activity against heterologous MenB strains and against serotypes A and C were also obtained:

	919		ORF46.1	
Strain	*Mixture*	*Hybrid*	*Mixture*	*Hybrid*
NGH38	1024	32000	-	16384
MC58	512	8192	-	512
BZ232	512	512	-	-
MenA (F6124)	512	32000	-	8192
MenC (C11)	>2048	>2048	-	-
MenC (BZ133)	>4096	64000	-	8192

The hybrid proteins with ΔG287 at the N-terminus are therefore immunologically superior to simple mixtures, with ΔG287-ORF46.1 being particularly effective, even against heterologous strains. ΔG287-ORF46.1K may be expressed in pET-24b.

The same hybrid proteins were made using New Zealand strain 394/98 rather than 2996:

ΔG287NZ-919

```
ATGGCTAGCCCCGATGTCAAGTCGGCGGACACGCTGTCAAAACCTGCCGCCCCTGTTGTTTCTGAAAAAGAGACAGAG
GCAAAGGAAGATGCGCCACAGGCAGGTTCTCAAGGACAGGGCGCGCCATCCGCACAAGGCGGTCAAGATATGGCGGCG
GTTTCGGAAGAAAATACAGGCAATGGCGGTGCGGCAGCAACGGACAAACCCAAAAATGAAGACGAGGGGGCGCAAAAT
GATATGCCGCAAAATGCCGCCGATACAGATAGTTTGACACCGAATCACACCCCGGCTTCGAATATGCCGGCCGGAAAT
ATGGAAAACCAAGCACCGGATGCCGGGGAATCGGAGCAGCCGGCAAACCAACCGGATATGGCAAATACGGCGGACGGA
ATGCAGGGTGACGATCCGTCGGCAGGCGGGGAAAATGCCGGCAATACGGCTGCCCAAGGTACAAATCAAGCCGAAAAC
AATCAAACCGCCGGTTCTCAAAATCCTGCCTCTTCAACCAATCCTAGCGCCACGAATAGCGGTGGTGATTTTGGAAGG
ACGAACGTGGGCAATTCTGTTGTGATTGACGGGCCGTCGCAAAATATAACGTTGACCCACTGTAAAGGCGATTCTTGT
AGTGGCAATAATTTCTTGGATGAAGAAGTACAGCTAAAATCAGAATTTGAAAAATTAAGTGATGCAGACAAAATAAGT
AATTACAAGAAAGATGGGAAGAATGACGGGAAGAATGATAAATTTGTCGGTTTGGTTGCCGATAGTGTGCAGATGAAG
GGAATCAATCAATATATTATCTTTTATAAACCTAAACCCACTTCATTTGCGCGATTTAGGCGTTCTGCACGGTCGAGG
CGGTCGCTTCCGGCCGAGATGCCGCTGATTCCCGTCAATCAGGCGGATACGCTGATTGTCGATGGGGAAGCGGTCAGC
CTGACGGGGCATTCCGGCAATATCTTCGCGCCCGAAGGGAATTACCGGTATCTGACTTACGGGGCGGAAAAATTGCCC
GGCGGATCGTATGCCCTCCGTGTTCAAGGCGAACCTTCAAAAGGCGAAATGCTCGCGGGCACGGCAGTGTACAACGGC
GAAGTGCTGCATTTTCATACGGAAAACGGCCGTCCGTCCCCGTCCAGAGGCAGGTTTGCCGCAAAAGTCGATTTCGGC
AGCAAATCTGTGGACGGCATTATCGACAGCGGCGATGGTTTGCATATGGGTACGCAAAAATTCAAAGCCGCCATCGAT
GGAAACGGCTTTAAGGGGACTTGGACGGAAAATGGCGGCGGGGATGTTTCCGGAAAGTTTTACGGCCCGGCCGGCGAG
```

```
GAAGTGGCGGGAAAATACAGCTATCGCCCAACAGATGCGGAAAAGGGCGGATTCGGCGTGTTTGCCGGCAAAAAAGAG
CAGGATGGATCCGGAGGAGGAGGATGCCAAAGCAAGAGCATCCAAACCTTTCCGCAACCCGACACATCCGTCATCAAC
GGCCCGGACCGGCCGGTCGGCATCCCCGACCCCGCCGGAACGACGGTCGGCGGCGGCGGGGCCGTCTATACCGTTGTA
CCGCACCTGTCCCTGCCCCACTGGGCGGCGCAGGATTTCGCCAAAAGCCTGCAATCCTTCCGCCTCGGCTGCGCCAAT
TTGAAAAACCGCCAAGGCTGGCAGGATGTGTGCGCCCAAGCCTTTCAAACCCCCGTCCATTCCTTTCAGGCAAAACAG
TTTTTTGAACGCTATTTCACGCCGTGGCAGGTTGCAGGCAACGGAAGCCTTGCCGGTACGGTTACCGGCTATTACGAG
CCGGTGCTGAAGGGCGACGACAGGCGGACGGCACAAGCCCGCTTCCCGATTTACGGTATTCCCGACGATTTTATCTCC
GTCCCCCTGCCTGCCGGTTTGCGGAGCGGAAAAGCCCTTGTCCGCATCAGGCAGACGGGAAAAAACAGCGGCACAATC
GACAATACCGGCGGCACACATACCGCCGACCTCTCCCGATTCCCCATCACCGCGCGCACAACGGCAATCAAAGGCAGG
TTTGAAGGAAGCCGCTTCCTCCCCTACCACACGCGCAACCAAATCAACGGCGGCGCGCTTGACGGCAAAGCCCCGATA
CTCGGTTACGCCGAAGACCCCGTCGAACTTTTTTTTATGCACATCCAAGGCTCGGGCCGTCTGAAAACCCCGTCCGGC
AAATACATCCGCATCGGCTATGCCGACAAAAACGAACATCCCTACGTTTCCATCGGACGCTATATGGCGGACAAAGGC
TACCTCAAGCTCGGGCAGACCTCGATGCAGGGCATCAAAGCCTATATGCGGCAAAATCCGCAACGCCTCGCCGAAGTT
TTGGGTCAAAACCCCAGCTATATCTTTTTCCGCGAGCTTGCCGGAAGCAGCAATGACGGTCCCGTCGGCGCACTGGGC
ACGCCGTTGATGGGGGAATATGCCGGCGCAGTCGACCGGCACTACATTACCTTGGGCGCGCCCTTATTTGTCGCCACC
GCCCATCCGGTTACCCGCAAAGCCCTCAACCGCCTGATTATGGCGCAGGATACCGGCAGCGCGATTAAAGGCGCGGTG
CGCGTGGATTATTTTTGGGGATACGGCGACGAAGCCGGCGAACTTGCCGGCAAACAGAAAACCACGGGTTACGTCTGG
CAGCTCCTACCCAACGGTATGAAGCCCGAATACCGCCCGTAAAAGCTT
```

```
  1  MASPDVKSAD TLSKPAAPVV SEKETEAKED APQAGSQGQG APSAQGGQDM
 51  AAVSEENTGN GGAAATDKPK NEDEGAQNDM PQNAADTDSL TPNHTPASNM
101  PAGNMENQAP DAGESEQPAN QPDMANTADG MQGDDPSAGG ENAGNTAAQG
151  TNQAENNQTA GSQNPASSTN PSATNSGGDF GRTNVGNSVV IDGPSQNITL
201  THCKGDSCSG NNFLDEEVQL KSEFEKLSDA DKISNYKKDG KNDGKNDKFV
251  GLVADSVQMK GINQYIIFYK PKPTSFARFR RSARSRRSLP AEMPLIPVNQ
301  ADTLIVDGEA VSLTGHSGNI FAPEGNYRYL TYGAEKLPGG SYALRVQGEP
351  SKGEMLAGTA VYNGEVLHFH TENGRPSPSR GRFAAKVDFG SKSVDGIIDS
401  GDGLHMGTQK FKAAIDGNGF KGTWTENGGG DVSGKFYGPA GEEVAGKYSY
451  RPTDAEKGGF GVFAGKKEQD GSGGGGCQSK SIQTFPQPDT SVINGPDRPV
501  GIPDPAGTTV GGGGAVYTVV PHLSLPHWAA QDFAKSLQSF RLGCANLKNR
551  QGWQDVCAQA FQTPVHSFQA KQFFERYFTP WQVAGNGSLA GTVTGYYEPV
601  LKGDDRRTAQ ARFPIYGIPD DFISVPLPAG LRSGKALVRI RQTGKNSGTI
651  DNTGGTHTAD LSRFPITART TAIKGRFEGS RFLPYHTRNQ INGGALDGKA
701  PILGYAEDPV ELFFMHIQGS GRLKTPSGKY IRIGYADKNE HPYVSIGRYM
751  ADKGYLKLGQ TSMQGIKAYM RQNPQRLAEV LGQNPSYIFF RELAGSSNDG
801  PVGALGTPLM GEYAGAVDRH YITLGAPLFV ATAHPVTRKA LNRLIMAQDT
851  GSAIKGAVRV DYFWGYGDEA GELAGKQKTT GYVWQLLPNG MKPEYRP*
```

<u>ΔG287NZ-953</u>

```
ATGGCTAGCCCCGATGTCAAGTCGGCGGACACGCTGTCAAAACCTGCCGCCCCTGTTGTTTCTGAAAAAGAGACAGAG
GCAAAGGAAGATGCGCCACAGGCAGGTTCTCAAGGACAGGGCGCGCCATCCGCACAAGGCGGTCAAGATATGGCGGCG
GTTTCGGAAGAAAATACAGGCAATGGCGGTGCGGCAGCAACGGACAAACCCAAAAATGAAGACGAGGGGGCGCAAAAT
GATATGCCGCAAAATGCCGCCGATACAGATAGTTTGACACCGAATCACACCCCGGCTTCGAATATGCCGGCCGGAAAT
ATGGAAAACCAAGCACCGGATGCCGGGGAATCGGAGCAGCCGGCAAACCAACCGGATATGGCAAATACGGCGGACGGA
ATGCAGGGTGACGATCCGTCGGCAGGCGGGGAAAATGCCGGCAATACGGCTGCCCAAGGTACAAATCAAGCCGAAAAC
AATCAAACCGCCGGTTCTCAAAATCCTGCCTCTTCAACCAATCCTAGCGCCACGAATAGCGGTGGTGATTTTGGAAGG
ACGAACGTGGGCAATTCTGTTGTGATTGACGGGCCGTCGCAAAATATAACGTTGACCCACTGTAAAGGCGATTCTTGT
AGTGGCAATAATTTCTTGGATGAAGAAGTACAGCTAAAATCAGAATTTGAAAAATTAAGTGATGCAGACAAAATAAGT
AATTACAAGAAAGATGGGAAGAATGACGGGAAGAATGATAAATTTGTCGGTTTGGTTGCCGATAGTGTGCAGATGAAG
GGAATCAATCAATATATTATCTTTTATAAACCTAAACCCACTTCATTTGCGCGATTTAGGCGTTCTGCACGGTCGAGG
CGGTCGCTTCCGGCCGAGATGCCGCTGATTCCCGTCAATCAGGCGGATACGCTGATTGTCGATGGGGAAGCGGTCAGC
CTGACGGGGCATTCCGGCAATATCTTCGCGCCCGAAGGGAATTACCGGTATCTGACTTACGGGGCGGAAAAATTGCCC
GGCGGATCGTATGCCCTCCGTGTTCAAGGCGAACCTTCAAAAGGCGAAATGCTCGCGGGCACGGCAGTGTACAACGGC
GAAGTGCTGCATTTTCATACGGAAAACGGCCGTCCGTCCCCGTCCAGAGGCAGGTTTGCCGCAAAAGTCGATTTCGGC
AGCAAATCTGTGGACGGCATTATCGACAGCGGCGATGGTTTGCATATGGGTACGCAAAAATTCAAAGCCGCCATCGAT
GGAAACGGCTTTAAGGGGACTTGGACGGAAAATGGCGGCGGGGATGTTTCCGGAAAGTTTTACGGCCCGGCCGGCGAG
GAAGTGGCGGGAAAATACAGCTATCGCCCAACAGATGCGGAAAAGGGCGGATTCGGCGTGTTTGCCGGCAAAAAAGAG
CAGGATGGATCCGGAGGAGGAGGAGCCACCTACAAAGTGGACGAATATCACGCCAACGCCCGTTTCGCCATCGACCAT
TTCAACACCAGCACCAACGTCGGCGGTTTTTACGGTCTGACCGGTTCCGTCGAGTTCGACCAAGCAAAACGCGACGGT
AAAATCGACATCACCATCCCCGTTGCCAACCTGCAAAGCGGTTCGCAACACTTTACCGACCACCTGAAATCAGCCGAC
ATCTTCGATGCCGCCCAATATCCGGACATCCGCTTTGTTTCCACCAAATTCAACTTCAACGGCAAAAAACTGGTTTCC
GTTGACGGCAACCTGACCATGCACGGCAAAACCGCCCCCGTCAAACTCAAAGCCGAAAAATTCAACTGCTACCAAAGC
CCGATGGCGAAAACCGAAGTTTGCGGCGGCGACTTCAGCACCACCATCGACCGCACCAAATGGGGCGTGGACTACCTC
GTTAACGTTGGTATGACCAAAAGCGTCCGCATCGACATCCAAATCGAGGCAGCCAAACAATAAAAGCTT
```

```
  1  MASPDVKSAD TLSKPAAPVV SEKETEAKED APQAGSQGQG APSAQGGQDM
```

51 AAVSEENTGN GGAAATDKPK NEDEGAQNDM PQNAADTDSL TPNHTPASNM
101 PAGNMENQAP DAGESEQPAN QPDMANTADG MQGDDPSAGG ENAGNTAAQG
151 TNQAENNQTA GSQNPASSTN PSATNSGGDF GRTNVGNSVV IDGPSQNITL
201 THCKGDSCSG NNFLDEEVQL KSEFEKLSDA DKISNYKKDG KNDGKNDKFV
251 GLVADSVQMK GINQYIIFYK PKPTSFARFR RSARSRRSLP AEMPLIPVNQ
301 ADTLIVDGEA VSLTGHSGNI FAPEGNYRYL TYGAEKLPGG SYALRVQGEP
351 SKGEMLAGTA VYNGEVLHFH TENGRPSPSR GRFAAKVDFG SKSVDGIIDS
401 GDGLHMGTQK FKAAIDGNGF KGTWTENGGG DVSGKFYGPA GEEVAGKYSY
451 RPTDAEKGGF GVFAGKKEQD GSGGGGATYK VDEYHANARF AIDHFNTSTN
501 VGGFYGLTGS VEFDQAKRDG KIDITIPVAN LQSGSQHFTD HLKSADIFDA
551 AQYPDIRFVS TKFNFNGKKL VSVDGNLTMH GKTAPVKLKA EKFNCYQSPM
601 AKTEVCGGDF STTIDRTKWG VDYLVNVGMT KSVRIDIQIE AAKQ*

<u>ΔG287NZ-961</u>

ATGGCTAGCCCCGATGTCAAGTCGGCGGACACGCTGTCAAAACCTGCCGCCCCTGTTGTTTCTGAAAAAGAGACAGAG
GCAAAGGAAGATGCGCCACAGGCAGGTTCTCAAGGACAGGGCGCGCCATCCGCACAAGGCGGTCAAGATATGGCGGCG
GTTTCGGAAGAAAATACAGGCAATGGCGGTGCGGCAGCAACGGACAAACCCAAAAATGAAGACGAGGGGGCGCAAAAT
GATATGCCGCAAAATGCCGCCGATACAGATAGTTTGACACCGAATCACACCCCGGCTTCGAATATGCCGGCCGGAAAT
ATGGAAAACCAAGCACCGGATGCCGGGGAATCGGAGCAGCCGGCAAACCAACCGGATATGGCAAATACGGCGGACGGA
ATGCAGGGTGACGATCCGTCGGCAGGCGGGGAAAATGCCGGCAATACGGCTGCCCAAGGTACAAATCAAGCCGAAAAC
AATCAAACCGCCGGTTCTCAAAATCCTGCCTCTTCAACCAATCCTAGCGCCACGAATAGCGGTGGTGATTTTGGAAGG
ACGAACGTGGGCAATTCTGTTGTGATTGACGGGCCGTCGCAAAATATAACGTTGACCCACTGTAAAGGCGATTCTTGT
AGTGGCAATAATTTCTTGGATGAAGAAGTACAGCTAAAATCAGAATTTGAAAAATTAAGTGATGCAGACAAAATAAGT
AATTACAAGAAAGATGGGAAGAATGACGGGAAGAATGATAAATTTGTCGGTTTGGTTGCCGATAGTGTGCAGATGAAG
GGAATCAATCAATATATTATCTTTTATAAACCTAAACCCACTTCATTTGCGCGATTTAGGCGTTCTGCACGGTCGAGG
CGGTCGCTTCCGGCCGAGATGCCGCTGATTCCCGTCAATCAGGCGGATACGCTGATTGTCGATGGGGAAGCGGTCAGC
CTGACGGGGCATTCCGGCAATATCTTCGCGCCCGAAGGGAATTACCGGTATCTGACTTACGGGGCGGAAAAATTGCCC
GGCGGATCGTATGCCCTCCGTGTTCAAGGCGAACCTTCAAAAGGCGAAATGCTCGCGGGCACGGCAGTGTACAACGGC
GAAGTGCTGCATTTTCATACGGAAAACGGCCGTCCGTCCCCGTCCAGAGGCAGGTTTGCCGCAAAAGTCGATTTCGGC
AGCAAATCTGTGGACGGCATTATCGACAGCGGCGATGGTTTGCATATGGGTACGCAAAAATTCAAAGCCGCCATCGAT
GGAAACGGCTTTAAGGGGACTTGGACGGAAAATGGCGGCGGGGATGTTTCCGGAAAGTTTTACGGCCCGGCCGGCGAG
GAAGTGGCGGGAAAATACAGCTATCGCCCAACAGATGCGGAAAAGGGCGGATTCGGCGTGTTTGCCGGCAAAAAAGAG
CAGGATGGATCCGGAGGAGGAGGAGCCACAAACGACGACGATGTTAAAAAAGCTGCCACTGTGGCCATTGCTGCTGCC
TACAACAATGGCCAAGAAATCAACGGTTTCAAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACC
AAAAAAGACGCAACTGCAGCCGATGTTGAAGCCGACGACTTTAAAGGTCTGGGTCTGAAAAAAGTCGTGACTAACCTG
ACCAAAACCGTCAATGAAAACAAACAAAACGTCGATGCCAAAGTAAAAGCTGCAGAATCTGAAATAGAAAAGTTAACA
ACCAAGTTAGCAGACACTGATGCCGCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAA
TTGGGAGAAAATATAACGACATTTGCTGAAGAGACTAAGACAAATATCGTAAAAATTGATGAAAAATTAGAAGCCGTG
GCTGATACCGTCGACAAGCATGCCGAAGCATTCAACGATATCGCCGATTCATTGGATGAAACCAACACTAAGGCAGAC
GAAGCCGTCAAAACCGCCAATGAAGCCAAACAGACGGCCGAAGAAACCAAACAAAACGTCGATGCCAAAGTAAAAGCT
GCAGAAACTGCAGCAGGCAAAGCCGAAGCTGCCGCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTCGCT
GCAAAAGTTACCGACATCAAAGCTGATATCGCTACGAACAAAGATAATATTGCTAAAAAAGCAAACAGTGCCGACGTG
TACACCAGAGAAGAGTCTGACAGCAAATTTGTCAGAATTGATGGTCTGAACGCTACTACCGAAAAATTGGACACACGC
TTGGCTTCTGCTGAAAAATCCATTGCCGATCACGATACTCGCCTGAACGGTTTGGATAAAACAGTGTCAGACCTGCGC
AAAGAAACCCGCCAAGGCCTTGCAGAACAAGCCGCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTCGGTTCAAT
GTAACGGCTGCAGTCGGCGGCTACAAATCCGAATCGGCAGTCGCCATCGGTACCGGCTTCCGCTTTACCGAAAACTTT
GCCGCCAAAGCAGGCGTGGCAGTCGGCACTTCGTCCGGTTCTTCCGCAGCCTACCATGTCGGCGTCAATTACGAGTGG
TAAAAGCTT

1 MASPDVKSAD TLSKPAAPVV SEKETEAKED APQAGSQGQG APSAQGGQDM
51 AAVSEENTGN GGAAATDKPK NEDEGAQNDM PQNAADTDSL TPNHTPASNM
101 PAGNMENQAP DAGESEQPAN QPDMANTADG MQGDDPSAGG ENAGNTAAQG
151 TNQAENNQTA GSQNPASSTN PSATNSGGDF GRTNVGNSVV IDGPSQNITL
201 THCKGDSCSG NNFLDEEVQL KSEFEKLSDA DKISNYKKDG KNDGKNDKFV
251 GLVADSVQMK GINQYIIFYK PKPTSFARFR RSARSRRSLP AEMPLIPVNQ
301 ADTLIVDGEA VSLTGHSGNI FAPEGNYRYL TYGAEKLPGG SYALRVQGEP
351 SKGEMLAGTA VYNGEVLHFH TENGRPSPSR GRFAAKVDFG SKSVDGIIDS
401 GDGLHMGTQK FKAAIDGNGF KGTWTENGGG DVSGKFYGPA GEEVAGKYSY
451 RPTDAEKGGF GVFAGKKEQD GSGGGGATND DDVKKAATVA IAAAYNNGQE
501 INGFKAGETI YDIDEDGTIT KKDATAADVE ADDFKGLGLK KVVTNLTKTV
551 NENKQNVDAK VKAAESEIEK LTTKLADTDA ALADTDAALD ATTNALNKLG
601 ENITTFAEET KTNIVKIDEK LEAVADTVDK HAEAFNDIAD SLDETNTKAD
651 EAVKTANEAK QTAEETKQNV DAKVKAAETA AGKAEAAAGT ANTAADKAEA
701 VAAKVTDIKA DIATNKDNIA KKANSADVYT REESDSKFVR IDGLNATTEK
751 LDTRLASAEK SIADHDTRLN GLDKTVSDLR KETRQGLAEQ AALSGLFQPY
801 NVGRFNVTAA VGGYKSESAV AIGTGFRFTE NFAAKAGVAV GTSSGSSAAY
851 HVGVNYEW*

Example 3 – hybrids of ΔG983

Protein 983 has the following sequence:

```
983                                              ➡ ΔG983
   1  MRTTPTFPTK TFKPTAMALA VATTLSACLG GGGGGTSAPD FNAGGTGIGS
  51  NSRATTAKSA AVSYAGIKNE MCKDRSMLCA GRDDVAVTDR DAKINAPPPN
 101  LHTGDFPNPN DAYKNLINLK PAIEAGYTGR GVEVGIVDTG ESVGSISFPE
 151  LYGRKEHGYN ENYKNYTAYM RKEAPEDGGG KDIEASFDDE AVIETEAKPT
 201  DIRHVKEIGH IDLVSHIIGG RSVDGRPAGG IAPDATLHIM NTNDETKNEM
 251  MVAAIRNAWV KLGERGVRIV NNSFGTTSRA GTADLFQIAN SEEQYRQALL
 301  DYSGGDKTDE GIRLMQQSDY GNLSYHIRNK NMLFIFSTGN DAQAQPNTYA
 351  LLPFYEKDAQ KGIITVAGVD RSGEKFKREM YGEPGTEPLE YGSNHCGITA
 401  MWCLSAPYEA SVRFTRTNPI QIAGTSFSAP IVTGTAALLL QKYPWMSNDN
 451  LRTTLLTTAQ DIGAVGVDSK FGWGLLDAGK AMNGPASFPF GDFTADTKGT
 501  SDIAYSFRND ISGTGGLIKK GGSQLQLHGN NTYTGKTIIE GGSLVLYGNN
 551  KSDMRVETKG ALIYNGAASG GSLNSDGIVY LADTDQSGAN ETVHIKGSLQ
 601  LDGKGTLYTR LGKLLKVDGT AIIGGKLYMS ARGKGAGYLN STGRRVPFLS
 651  AAKIGQDYSF FTNIETDGGL LASLDSVEKT AGSEGDTLSY YVRRGNAART
 701  ASAAAHSAPA GLKHAVEQGG SNLENLMVEL DASESSATPE TVETAAADRT
 751  DMPGIRPYGA TFRAAAAVQH ANAADGVRIF NSLAATVYAD STAAHADMQG
 801  RRLKAVSDGL DHNGTGLRVI AQTQQDGGTW EQGGVEGKMR GSTQTVGIAA
 851  KTGENTTAAA TLGMGRSTWS ENSANAKTDS ISLFAGIRHD AGDIGYLKGL
 901  FSYGRYKNSI SRSTGADEHA EGSVNGTLMQ LGALGGVNVP FAATGDLTVE
 951  GGLRYDLLKQ DAFAEKGSAL GWSGNSLTEG TLVGLAGLKL SQPLSDKAVL
1001  FATAGVERDL NGRDYTVTGG FTGATAATGK TGARNMPHTR LVAGLGADVE
1051  FGNGWNGLAR YSYAGSKQYG NHSGRVGVGY RF*
```

ΔG983 thus has the following basic sequence:

```
                                          TSAPD FNAGGTGIGS
NSRATTAKSA AVSYAGIKNE MCKDRSMLCA GRDDVAVTDR DAKINAPPPN
LHTGDFPNPN DAYKNLINLK PAIEAGYTGR GVEVGIVDTG ESVGSISFPE
LYGRKEHGYN ENYKNYTAYM RKEAPEDGGG KDIEASFDDE AVIETEAKPT
DIRHVKEIGH IDLVSHIIGG RSVDGRPAGG IAPDATLHIM NTNDETKNEM
MVAAIRNAWV KLGERGVRIV NNSFGTTSRA GTADLFQIAN SEEQYRQALL
DYSGGDKTDE GIRLMQQSDY GNLSYHIRNK NMLFIFSTGN DAQAQPNTYA
LLPFYEKDAQ KGIITVAGVD RSGEKFKREM YGEPGTEPLE YGSNHCGITA
MWCLSAPYEA SVRFTRTNPI QIAGTSFSAP IVTGTAALLL QKYPWMSNDN
LRTTLLTTAQ DIGAVGVDSK FGWGLLDAGK AMNGPASFPF GDFTADTKGT
SDIAYSFRND ISGTGGLIKK GGSQLQLHGN NTYTGKTIIE GGSLVLYGNN
KSDMRVETKG ALIYNGAASG GSLNSDGIVY LADTDQSGAN ETVHIKGSLQ
LDGKGTLYTR LGKLLKVDGT AIIGGKLYMS ARGKGAGYLN STGRRVPFLS
AAKIGQDYSF FTNIETDGGL LASLDSVEKT AGSEGDTLSY YVRRGNAART
ASAAAHSAPA GLKHAVEQGG SNLENLMVEL DASESSATPE TVETAAADRT
DMPGIRPYGA TFRAAAAVQH ANAADGVRIF NSLAATVYAD STAAHADMQG
RRLKAVSDGL DHNGTGLRVI AQTQQDGGTW EQGGVEGKMR GSTQTVGIAA
KTGENTTAAA TLGMGRSTWS ENSANAKTDS ISLFAGIRHD AGDIGYLKGL
FSYGRYKNSI SRSTGADEHA EGSVNGTLMQ LGALGGVNVP FAATGDLTVE
GGLRYDLLKQ DAFAEKGSAL GWSGNSLTEG TLVGLAGLKL SQPLSDKAVL
FATAGVERDL NGRDYTVTGG FTGATAATGK TGARNMPHTR LVAGLGADVE
 FGNGWNGLAR YSYAGSKQYG NHSGRVGVGY RF*
```

ΔG983 was expressed as a hybrid, with ORF46.1, 741, 961 or 961c at its C-terminus:

ΔG983-ORF46.1

```
ATGACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA
GCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCG
GTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCA
TACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGAC
ACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAA
AACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAG
GCCGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCCCAT
```

ATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACG
AATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGC
ATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAG
TACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTAC
GGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCC
AACACATATGCCCTATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGT
GGAGAAAAGTTCAAACGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCATTGCGGAATT
ACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGA
ACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAAC
CTGCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTG
GATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACATCCGAT
ATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCAC
GGCAACAACACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATG
CGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGCGACGGCATTGTCTAT
CTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACG
CTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGC
GGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGGCAGGATTAT
TCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAA
GGCGACACGCTGTCCTATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCC
GGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCC
GCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTC
CGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTAT
GCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGC
ACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGC
GGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGACGC
AGCACATGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGC
GATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAA
CATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACG
GGAGATTTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTG
GGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGCGAT
AAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACC
GGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGAT
GTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGA
CGAGTCGGCGTAGGCTACCGGTTCCTCGACGGTGGCGGAGGCACTGGATCCTCAGATTTGGCAAACGATTCTTTTATC
CGGCAGGTTCTCGACCGTCAGCATTTCGAACCCGACGGAAATACCACCTATTCGGCAGCAGGGGGGAACTTGCCGAG
CGCAGCGGCCATATCGGATTGGGAAAAATACAAAGCCATCAGTTGGGCAACCTGATGATTCAACAGGCGGCCATTAAA
GGAAATATCGGCTACATTGTCCGCTTTTCCGATCACGGGCACGAAGTCCATTCCCCCTTCGACAACCATGCCTCACAT
TCCGATTCTGATGAAGCCGGTAGTCCCGTTGACGGATTTAGCCTTTACCGCATCCATTGGGACGGATACGAACACCAT
CCCGCCGACGGCTATGACGGGCCACAGGGCGGCGGCTATCCCGCTCCCAAAGGCGCGAGGGATATATACAGCTACGAC
ATAAAAGGCGTTGCCCAAAATATCCGCCTCAACCTGACCGACAACCGCAGCACCGGACAACGGCTTGCCGACCGTTTC
CACAATGCCGGTAGTATGCTGACGCAAGGAGTAGGCGACGGATTCAAACGCGCCACCCGATACAGCCCCGAGCTGGAC
AGATCGGGCAATGCCGCCGAAGCCTTCAACGGCACTGCAGATATCGTTAAAAACATCATCGGCGCGGCAGGAGAAATT
GTCGGCGCAGGCGATGCCGTGCAGGGCATAAGCGAAGGCTCAAACATTGCTGTCATGCACGGCTTGGGTCTGCTTTCC
ACCGAAAACAAGATGGCGCGCATCAACGATTTGGCAGATATGGCGCAACTCAAAGACTATGCCGCAGCAGCCATCCGC
GATTGGGCAGTCCAAAACCCCAATGCCGCACAAGGCATAGAAGCCGTCAGCAATATCTTTATGGCAGCCATCCCCATC
AAAGGGATTGGAGCTGTTCGGGGAAAATACGGCTTGGGCGGCATCACGGCACATCCTATCAAGCGGTCGCAGATGGGC
GCGATCGCATTGCCGAAAGGGAAATCCGCCGTCAGCGACAATTTTGCCGATGCGGCATACGCCAAATACCCGTCCCCT
TACCATTCCCGAAATATCCGTTCAAACTTGGAGCAGCGTTACGGCAAAGAAAACATCACCTCCTCAACCGTGCCGCCG
TCAAACGGCAAAAATGTCAAACTGGCAGACCAACGCCACCCGAAGACAGGCGTACCGTTTGACGGTAAAGGGTTTCCG
AATTTTGAGAAGCACGTGAAATATGATACGCTCGAGCACCACCACCACCACCACTGA

1	MTSAPDFNAG	GTGIGSNSRA	TTAKSAAVSY	AGIKNEMCKD	RSMLCAGRDD
51	VAVTDRDAKI	NAPPPNLHTG	DFPNPNDAYK	NLINLKPAIE	AGYTGRGVEV
101	GIVDTGESVG	SISFPELYGR	KEHGYNENYK	NYTAYMRKEA	PEDGGGKDIE
151	ASFDDEAVIE	TEAKPTDIRH	VKEIGHIDLV	SHIIGGRSVD	GRPAGGIAPD
201	ATLHIMNTND	ETKNEMMVAA	IRNAWVKLGE	RGVRIVNNSF	GTTSRAGTAD
251	LFQIANSEEQ	YRQALLDYSG	GDKTDEGIRL	MQQSDYGNLS	YHIRNKNMLF
301	IFSTGNDAQA	QPNTYALLPF	YEKDAQKGII	TVAGVDRSGE	KFKREMYGEP
351	GTEPLEYGSN	HCGITAMWCL	SAPYEASVRF	TRTNPIQIAG	TSFSAPIVTG
401	TAALLLQKYP	WMSNDNLRTT	LLTTAQDIGA	VGVDSKFGWG	LLDAGKAMNG
451	PASFPFGDFT	ADTKGTSDIA	YSFRNDISGT	GGLIKKGGSQ	LQLHGNNTYT
501	GKTIIEGGSL	VLYGNNKSDM	RVETKGALIY	NGAASGGSLN	SDGIVYLADT
551	DQSGANETVH	IKGSLQLDGK	GTLYTRLGKL	LKVDGTAIIG	GKLYMSARGK
601	GAGYLNSTGR	RVPFLSAAKI	GQDYSFFTNI	ETDGGLLASL	DSVEKTAGSE
651	GDTLSYYVRR	GNAARTASAA	AHSAPAGLKH	AVEQGGSNLE	NLMVELDASE
701	SSATPETVET	AAADRTDMPG	IRPYGATFRA	AAAVQHANAA	DGVRIFNSLA
751	ATVYADSTAA	HADMQGRRLK	AVSDGLDHNG	TGLRVIAQTQ	QDGGTWEQGG

```
801   VEGKMRGSTQ TVGIAAKTGE NTTAAATLGM GRSTWSENSA NAKTDSISLF
851   AGIRHDAGDI GYLKGLFSYG RYKNSISRST GADEHAEGSV NGTLMQLGAL
901   GGVNVPFAAT GDLTVEGGLR YDLLKQDAFA EKGSALGWSG NSLTEGTLVG
951   LAGLKLSQPL SDKAVLFATA GVERDLNGRD YTVTGGFTGA TAATGKTGAR
1001  NMPHTRLVAG LGADVEFGNG WNGLARYSYA GSKQYGNHSG RVGVGYRFLD
1051  GGGGTGSSDL ANDSFIRQVL DRQHFEPDGK YHLFGSRGEL AERSGHIGLG
1101  KIQSHQLGNL MIQQAAIKGN IGYIVRFSDH GHEVHSPFDN HASHSDSDEA
1151  GSPVDGFSLY RIHWDGYEHH PADGYDGPQG GGYPAPKGAR DIYSYDIKGV
1201  AQNIRLNLTD NRSTGQRLAD RFHNAGSMLT QGVGDGFKRA TRYSPELDRS
1251  GNAAEAFNGT ADIVKNIIGA AGEIVGAGDA VQGISEGSNI AVMHGLGLLS
1301  TENKMARIND LADMAQLKDY AAAAIRDWAV QNPNAAQGIE AVSNIFMAAI
1351  PIKGIGAVRG KYGLGGITAH PIKRSQMGAI ALPKGKSAVS DNFADAAYAK
1401  YPSPYHSRNI RSNLEQRYGK ENITSSTVPP SNGKNVKLAD QRHPKTGVPF
1451  DGKGFPNFEK HVKYDTLEHH HHHH*
```

ΔG983-741

ATGACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA
GCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGATGACGTTGCG
GTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCA
TACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGAC
ACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAA
AACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAG
GCCGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCCCAT
ATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACG
AATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGC
ATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAG
TACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTAC
GGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCC
AACACATATGCCCTATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGT
GGAGAAAAGTTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCATTGCGGAATT
ACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGA
ACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAAC
CTGCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTG
GATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACATCCGAT
ATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCAC
GGCAACAACACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATG
CGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGCGACGGCATTGTCTAT
CTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACG
CTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGC
GGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGGCAGGATTAT
TCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAA
GGCGACACGCTGTCCTATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCC
GGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCC
GCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTC
CGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTAT
GCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGC
ACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGC
GGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGACGC
AGCACATGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGC
GATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAA
CATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACG
GGAGATTTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTG
GGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGCGAT
AAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACC
GGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGAT
GTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGA
CGAGTCGGCGTAGGCTACCGGTTCCTCGAGGGATCCGGAGGGGGTGGTGTCGCCGCCGACATCGGTGCGGGGCTTGCC
GATGCACTAACCGCACCGCTCGACCATAAAGACAAAGGTTTGCAGTCTTTGACGCTGGATCAGTCCGTCAGGAAAAAC
GAGAAACTGAAGCTGGCGGCACAAGGTGCGGAAAAAACTTATGGAAACGGTGACAGCCTCAATACGGGCAAATTGAAG
AACGACAAGGTCAGCCGTTTCGACTTTATCCGCCAAATCGAAGTGGACGGGCAGCTCATTACCTTGGAGAGTGGAGAG
TTCCAAGTATACAAACAAAGCCATTCCGCCTTAACCGCCTTTCAGACCGAGCAAATACAAGATTCGGAGCATTCCGGG
AAGATGGTTGCGAAACGCCAGTTCAGAATCGGCGACATAGCGGGCGAACATACATCTTTTGACAAGCTTCCCGAAGGC
GGCAGGGCGACATATCGCGGGACGGCGTTCGGTTCAGACGATGCCGGCGGAAAACTGACCTACACCATAGATTTCGCC
GCCAAGCAGGGAAACGGCAAAATCGAACATTTGAAATCGCCAGAACTCAATGTCGACCTGGCCGCCGCCGATATCAAG
CCGGATGGAAAACGCCATGCCGTCATCAGCGGTTCCGTCCTTTACAACCAAGCCGAGAAAGGCAGTTACTCCCTCGGT

ATCTTTGGCGGAAAAGCCCAGGAAGTTGCCGGCAGCGCGGAAGTGAAAACCGTAAACGGCATACGCCATATCGGCCTT
GCCGCCAAGCAACTCGAGCACCACCACCACCACCACTGA

```
   1  MTSAPDFNAG GTGIGSNSRA TTAKSAAVSY AGIKNEMCKD RSMLCAGRDD
  51  VAVTDRDAKI NAPPPNLHTG DFPNPNDAYK NLINLKPAIE AGYTGRGVEV
 101  GIVDTGESVG SISFPELYGR KEHGYNENYK NYTAYMRKEA PEDGGGKDIE
 151  ASFDDEAVIE TEAKPTDIRH VKEIGHIDLV SHIIGGRSVD GRPAGGIAPD
 201  ATLHIMNTND ETKNEMMVAA IRNAWVKLGE RGVRIVNNSF GTTSRAGTAD
 251  LFQIANSEEQ YRQALLDYSG GDKTDEGIRL MQQSDYGNLS YHIRNKNMLF
 301  IFSTGNDAQA QPNTYALLPF YEKDAQKGII TVAGVDRSGE KFKREMYGEP
 351  GTEPLEYGSN HCGITAMWCL SAPYEASVRF TRTNPIQIAG TSFSAPIVTG
 401  TAALLLQKYP WMSNDNLRTT LLTTAQDIGA VGVDSKFGWG LLDAGKAMNG
 451  PASFPFGDFT ADTKGTSDIA YSFRNDISGT GGLIKKGGSQ LQLHGNNTYT
 501  GKTIIEGGSL VLYGNNKSDM RVETKGALIY NGAASGGSLN SDGIVYLADT
 551  DQSGANETVH IKGSLQLDGK GTLYTRLGKL LKVDGTAIIG GKLYMSARGK
 601  GAGYLNSTGR RVPFLSAAKI GQDYSFFTNI ETDGGLLASL DSVEKTAGSE
 651  GDTLSYYVRR GNAARTASAA AHSAPAGLKH AVEQGGSNLE NLMVELDASE
 701  SSATPETVET AAADRTDMPG IRPYGATFRA AAVQHANAA DGVRIFNSLA
 751  ATVYADSTAA HADMQGRRLK AVSDGLDHNG TGLRVIAQTQ QDGGTWEQGG
 801  VEGKMRGSTQ TVGIAAKTGE NTTAAATLGM GRSTWSENSA NAKTDSISLF
 851  AGIRHDAGDI GYLKGLFSYG RYKNSISRST GADEHAEGSV NGTLMQLGAL
 901  GGVNVPFAAT GDLTVEGGLR YDLLKQDAFA EKGSALGWSG NSLTEGTLVG
 951  LAGLKLSQPL SDKAVLFATA GVERDLNGRD YTVTGGFTGA TAATGKTGAR
1001  NMPHTRLVAG LGADVEFGNG WNGLARYSYA GSKQYGNHSG RVGVGYRFLE
1051  GSGGGGVAAD IGAGLADALT APLDHKDKGL QSLTLDQSVR KNEKLKLAAQ
1101  GAEKTYGNGD SLNTGKLKND KVSRFDFIRQ IEVDGQLITL ESGEFQVYKQ
1151  SHSALTAFQT EQIQDSEHSG KMVAKRQFRI GDIAGEHTSF DKLPEGGRAT
1201  YRGTAFGSDD AGGKLTYTID FAAKQGNGKI EHLKSPELNV DLAAADIKPD
1251  GKRHAVISGS VLYNQAEKGS YSLGIFGGKA QEVAGSAEVK TVNGIRHIGL
1301  AAKQLEHHHH HH*
```

ΔG983-961

ATGACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA
GCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCG
GTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCA
TACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGAC
ACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAA
AACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAG
GCCGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCCCAT
ATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACG
AATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGC
ATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAG
TACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTAC
GGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCC
AACACATATGCCCTATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGT
GGAGAAAAGTTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCATTGCGGAATT
ACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGA
ACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAAC
CTGCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTG
GATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACATCCGAT
ATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCAC
GGCAACAACACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATG
CGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGCGACGGCATTGTCTAT
CTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACG
CTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGC
GGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGGCAGGATTAT
TCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAA
GGCGACACGCTGTCCTATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCC
GGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCC

```
CATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACG
GGAGATTTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTG
GGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGCGAT
AAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACC
GGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGAT
GTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGA
CGAGTCGGCGTAGGCTACCGGTTCCTCGAGGGTGGCGGAGGCACTGGATCCGCCACAAACGACGACGATGTTAAAAAA
GCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATCAACGGTTTCAAAGCTGGAGAGACCATCTAC
GACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCCGATGTTGAAGCCGACGACTTTAAAGGTCTG
GGTCTGAAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAACAAACAAAACGTCGATGCCAAAGTAAAAGCT
GCAGAATCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGATGCCGCTTTAGCAGATACTGATGCCGCTCTG
GATGCAACCACCAACGCCTTGAATAAATTGGGAGAAAATATAACGACATTTGCTGAAGAGACTAAGACAAATATCGTA
AAAATTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCATGCCGAAGCATTCAACGATATCGCCGATTCA
TTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAATGAAGCCAAACAGACGGCCGAAGAAACCAAA
CAAAACGTCGATGCCAAAGTAAAAGCTGCAGAAACTGCAGCAGGCAAAGCCGAAGCTGCCGCTGGCACAGCTAATACT
GCAGCCGACAAGGCCGAAGCTGTCGCTGCAAAAGTTACCGACATCAAAGCTGATATCGCTACGAACAAAGATAATATT
GCTAAAAAAGCAAACAGTGCCGACGTGTACACCAGAGAAGAGTCTGACAGCAAATTTGTCAGAATTGATGGTCTGAAC
GCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCCATTGCCGATCACGATACTCGCCTGAACGGT
TTGGATAAAACAGTGTCAGACCTGCGCAAAGAAACCCGCCAAGGCCTTGCAGAACAAGCCGCGCTCTCCGGTCTGTTC
CAACCTTACAACGTGGGTCGGTTCAATGTAACGGCTGCAGTCGGCGGCTACAAATCCGAATCGGCAGTCGCCATCGGT
ACCGGCTTCCGCTTTACCGAAAACTTTGCCGCCAAAGCAGGCGTGGCAGTCGGCACTTCGTCCGGTTCTTCCGCAGCC
TACCATGTCGGCGTCAATTACGAGTGGCTCGAGCACCACCACCACCACCACTGA
```

```
   1  MTSAPDFNAG GTGIGSNSRA TTAKSAAVSY AGIKNEMCKD RSMLCAGRDD
  51  VAVTDRDAKI NAPPPNLHTG DFPNPNDAYK NLINLKPAIE AGYTGRGVEV
 101  GIVDTGESVG SISFPELYGR KEHGYNENYK NYTAYMRKEA PEDGGGKDIE
 151  ASFDDEAVIE TEAKPTDIRH VKEIGHIDLV SHIIGGRSVD GRPAGGIAPD
 201  ATLHIMNTND ETKNEMMVAA IRNAWVKLGE RGVRIVNNSF GTTSRAGTAD
 251  LFQIANSEEQ YRQALLDYSG GDKTDEGIRL MQQSDYGNLS YHIRNKNMLF
 301  IFSTGNDAQA QPNTYALLPF YEKDAQKGII TVAGVDRSGE KFKREMYGEP
 351  GTEPLEYGSN HCGITAMWCL SAPYEASVRF TRTNPIQIAG TSFSAPIVTG
 401  TAALLLQKYP WMSNDNLRTT LLTTAQDIGA VGVDSKFGWG LLDAGKAMNG
 451  PASFPFGDFT ADTKGTSDIA YSFRNDISGT GGLIKKGGSQ LQLHGNNTYT
 501  GKTIIEGGSL VLYGNNKSDM RVETKGALIY NGAASGGSLN SDGIVYLADT
 551  DQSGANETVH IKGSLQLDGK GTLYTRLGKL LKVDGTAIIG GKLYMSARGK
 601  GAGYLNSTGR RVPFLSAAKI GQDYSFFTNI ETDGGLLASL DSVEKTAGSE
 651  GDTLSYYVRR GNAARTASAA AHSAPAGLKH AVEQGGSNLE NLMVELDASE
 701  SSATPETVET AAADRTDMPG IRPYGATFRA AAVQHANAA DGVRIFNSLA
 751  ATVYADSTAA HADMQGRRLK AVSDGLDHNG TGLRVIAQTQ QDGGTWEQGG
 801  VEGKMRGSTQ TVGIAAKTGE NTTAAATLGM GRSTWSENSA NAKTDSISLF
 851  AGIRHDAGDI GYLKGLFSYG RYKNSISRST GADEHAEGSV NGTLMQLGAL
 901  GGVNVPFAAT GDLTVEGGLR YDLLKQDAFA EKGSALGWSG NSLTEGTLVG
 951  LAGLKLSQPL SDKAVLFATA GVERDLNGRD YTVTGGFTGA TAATGKTGAR
1001  NMPHTRLVAG LGADVEFGNG WNGLARYSYA GSKQYGNHSG RVGVGYRFLE
1051  GGGGTGSATN DDDVKKAATV AIAAAYNNGQ EINGFKAGET IYDIDEDGTI
1101  TKKDATAADV EADDFKGLGL KKVVTNLTKT VNENKQNVDA KVKAAESEIE
1151  KLTTKLADTD AALADTDAAL DATTNALNKL GENITTFAEE TKTNIVKIDE
1201  KLEAVADTVD KHAEAFNDIA DSLDETNTKA DEAVKTANEA KQTAEETKQN
1251  VDAKVKAAET AAGKAEAAAG TANTAADKAE AVAAKVTDIK ADIATNKDNI
1301  AKKANSADVY TREESDSKFV RIDGLNATTE KLDTRLASAE KSIADHDTRL
1351  NGLDKTVSDL RKETRQGLAE QAALSGLFQP YNVGRFNVTA AVGGYKSESA
1401  VAIGTGFRFT ENFAAKAGVA VGTSSGSSAA YHVGVNYEWL EHHHHHH*
```

ΔG983-961c

```
ATGACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA
GCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCG
GTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCA
TACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGAC
ACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAA
AACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAG
GCCGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCCCAT
ATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACG
AATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGC
ATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAG
TACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTAC
GGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCC
```

AACACATATGCCCTATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGT
GGAGAAAAGTTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCATTGCGGAATT
ACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGA
ACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCAGAAATACCCGTGGATGAGCAACGACAAC
CTGCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTG
GATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACATCCGAT
ATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCAC
GGCAACAACACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATG
CGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGCGACGGCATTGTCTAT
CTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACG
CTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGC
GGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGGCAGGATTAT
TCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAA
GGCGACACGCTGTCCTATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCC
GGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCC
GCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTC
CGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTAT
GCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGC
ACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGC
GGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGC
AGCACATGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGC
GATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAA
CATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACG
GGAGATTTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTG
GGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGCGAT
AAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACC
GGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGAT
GTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGA
CGAGTCGGCGTAGGCTACCGGTTCCTCGAGGGTGGCGGAGGCACTGGATCCGCCACAAACGACGACGATGTTAAAAAA
GCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATCAACGGTTTCAAAGCTGGAGAGACCATCTAC
GACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCCGATGTTGAAGCCGACGACTTTAAAGGTCTG
GGTCTGAAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAACAAACAAAACGTCGATGCCAAAGTAAAAGCT
GCAGAATCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGATGCCGCTTTAGCAGATACTGATGCCGCTCTG
GATGCAACCACCAACGCCTTGAATAAATTGGGAGAAAATATAACGACATTTGCTGAAGAGACTAAGACAAATATCGTA
AAAATTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCATGCCGAAGCATTCAACGATATCGCCGATTCA
TTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAATGAAGCCAAACAGACGGCCGAAGAAACCAAA
CAAAACGTCGATGCCAAAGTAAAAGCTGCAGAAACTGCAGCAGGCAAAGCCGAAGCTGCCGCTGGCACAGCTAATACT
GCAGCCGACAAGGCCGAAGCTGTCGCTGCAAAAGTTACCGACATCAAAGCTGATATCGCTACGAACAAAGATAATATT
GCTAAAAAAGCAAACAGTGCCGACGTGTACACCAGAGAAGAGTCTGACAGCAAATTTGTCAGAATTGATGGTCTGAAC
GCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCCATTGCCGATCACGATACTCGCCTGAACGGT
TTGGATAAAACAGTGTCAGACCTGCGCAAAGAAACCCGCCAAGGCCTTGCAGAACAAGCCGCGCTCTCCGGTCTGTTC
CAACCTTACAACGTGGGTCTCGAGCACCACCACCACCACCACTGA

```
   1  MTSAPDFNAG GTGIGSNSRA TTAKSAAVSY AGIKNEMCKD RSMLCAGRDD
  51  VAVTDRDAKI NAPPPNLHTG DFPNPNDAYK NLINLKPAIE AGYTGRGVEV
 101  GIVDTGESVG SISFPELYGR KEHGYNENYK NYTAYMRKEA PEDGGGKDIE
 151  ASFDDEAVIE TEAKPTDIRH VKEIGHIDLV SHIIGGRSVD GRPAGGIAPD
 201  ATLHIMNTND ETKNEMMVAA IRNAWVKLGE RGVRIVNNSF GTTSRAGTAD
 251  LFQIANSEEQ YRQALLDYSG GDKTDEGIRL MQQSDYGNLS YHIRNKNMLF
 301  IFSTGNDAQA QPNTYALLPF YEKDAQKGII TVAGVDRSGE KFKREMYGEP
 351  GTEPLEYGSN HCGITAMWCL SAPYEASVRF TRTNPIQIAG TSFSAPIVTG
 401  TAALLLQKYP WMSNDNLRTT LLTTAQDIGA VGVDSKFGWG LLDAGKAMNG
 451  PASFPFGDFT ADTKGTSDIA YSFRNDISGT GGLIKKGGSQ LQLHGNNTYT
 501  GKTIIEGGSL VLYGNNKSDM RVETKGALIY NGAASGGSLN SDGIVYLADT
 551  DQSGANETVH IKGSLQLDGK GTLYTRLGKL LKVDGTAIIG GKLYMSARGK
 601  GAGYLNSTGR RVPFLSAAKI GQDYSFFTNI ETDGGLLASL DSVEKTAGSE
 651  GDTLSYYVRR GNAARTASAA AHSAPAGLKH AVEQGGSNLE NLMVELDASE
 701  SSATPETVET AAADRTDMPG IRPYGATFRA AAVQHANAA DGVRIFNSLA
 751  ATVYADSTAA HADMQGRRLK AVSDGLDHNG TGLRVIAQTQ QDGGTWEQGG
 801  VEGKMRGSTQ TVGIAAKTGE NTTAAATLGM GRSTWSENSA NAKTDSISLF
 851  AGIRHDAGDI GYLKGLFSYG RYKNSISRST GADEHAEGSV NGTLMQLGAL
 901  GGVNVPFAAT GDLTVEGGLR YDLLKQDAFA EKGSALGWSG NSLTEGTLVG
 951  LAGLKLSQPL SDKAVLFATA GVERDLNGRD YTVTGGFTGA TAATGKTGAR
1001  NMPHTRLVAG LGADVEFGNG WNGLARYSYA GSKQYGNHSG RVGVGYRFLE
1051  GGGGTGSATN DDDVKKAATV AIAAAYNNGQ EINGFKAGET IYDIDEDGTI
1101  TKKDATAADV EADDFKGLGL KKVVTNLTKT VNENKQNVDA KVKAAESEIE
1151  KLTTKLADTD AALADTDAAL DATTNALNKL GENITTFAEE TKTNIVKIDE
```

```
1201   KLEAVADTVD KHAEAFNDIA DSLDETNTKA DEAVKTANEA KQTAEETKQN
1251   VDAKVKAAET AAGKAEAAAG TANTAADKAE AVAAKVTDIK ADIATNKDNI
1301   AKKANSADVY TREESDSKFV RIDGLNATTE KLDTRLASAE KSIADHDTRL
1351   NGLDKTVSDL RKETRQGLAE QAALSGLFQP YNVGLEHHHH HH*
```

Example 4 – hybrids of ΔG741

Protein 741 has the following sequence:

```
  1  VNRTAFCCLS LTTALILTAC SSGGGGVAAD IGAGLADALT APLDHKDKGL
 51  QSLTLDQSVR KNEKLKLAAQ GAEKTYGNGD SLNTGKLKND KVSRFDFIRQ
101  IEVDGQLITL ESGEFQVYKQ SHSALTAFQT EQIQDSEHSG KMVAKRQFRI
151  GDIAGEHTSF DKLPEGGRAT YRGTAFGSDD AGGKLTYTID FAAKQGNGKI
201  EHLKSPELNV DLAAADIKPD GKRHAVISGS VLYNQAEKGS YSLGIFGGKA
251  QEVAGSAEVK TVNGIRHIGL AAKQ*
```

ΔG741 thus has the following basic sequence:

```
                                 VAAD IGAGLADALT APLDHKDKGL
QSLTLDQSVR KNEKLKLAAQ GAEKTYGNGD SLNTGKLKND KVSRFDFIRQ
IEVDGQLITL ESGEFQVYKQ SHSALTAFQT EQIQDSEHSG KMVAKRQFRI
GDIAGEHTSF DKLPEGGRAT YRGTAFGSDD AGGKLTYTID FAAKQGNGKI
EHLKSPELNV DLAAADIKPD GKRHAVISGS VLYNQAEKGS YSLGIFGGKA
QEVAGSAEVK TVNGIRHIGL AAKQ*
```

ΔG741 was fused directly in-frame upstream of proteins 961, 961c, 983 and ORF46.1:

ΔG741–961

```
ATGGTCGCCGCCGACATCGGTGCGGGGCTTGCCGATGCACTAACCGCACCGCTCGACCATAAAGACAAAGGTTTGCAG
TCTTTGACGCTGGATCAGTCCGTCAGGAAAAACGAGAAACTGAAGCTGGCGGCACAAGGTGCGGAAAAAACTTATGGA
AACGGTGACAGCCTCAATACGGGCAAATTGAAGAACGACAAGGTCAGCCGTTTCGACTTTATCCGCCAAATCGAAGTG
GACGGGCAGCTCATTACCTTGGAGAGTGGAGAGTTCCAAGTATACAAACAAAGCCATTCCGCCTTAACCGCCTTTCAG
ACCGAGCAAATACAAGATTCGGAGCATTCCGGGAAGATGGTTGCGAAACGCCAGTTCAGAATCGGCGACATAGCGGGC
GAACATACATCTTTTGACAAGCTTCCCGAAGGCGGCAGGGCGACATATCGCGGGACGGCGTTCGGTTCAGACGATGCC
GGCGGAAAACTGACCTACACCATAGATTTCGCCGCCAAGCAGGGAAACGGCAAAATCGAACATTTGAAATCGCCAGAA
CTCAATGTCGACCTGGCCGCCGCCGATATCAAGCCGGATGGAAAACGCCATGCCGTCATCAGCGGTTCCGTCCTTTAC
AACCAAGCCGAGAAAGGCAGTTACTCCCTCGGTATCTTTGGCGGAAAAGCCCAGGAAGTTGCCGGCAGCGCGGAAGTG
AAAACCGTAAACGGCATACGCCATATCGGCCTTGCCGCCAAGCAACTCGAGGGTGGCGGAGGCACTGGATCCGCCACA
AACGACGACGATGTTAAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATCAACGGTTTC
AAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCCGATGTTGAA
GCCGACGACTTTAAAGGTCTGGGTCTGAAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAACAAACAAAAC
GTCGATGCCAAAGTAAAAGCTGCAGAATCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGATGCCGCTTTA
GCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAATTGGGAGAAAATATAACGACATTTGCTGAA
GAGACTAAGACAAATATCGTAAAAATTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCATGCCGAAGCA
TTCAACGATATCGCCGATTCATTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAATGAAGCCAAA
CAGACGGCCGAAGAAACCAAACAAAACGTCGATGCCAAAGTAAAAGCTGCAGAAACTGCAGCAGGCAAAGCCGAAGCT
GCCGCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTCGCTGCAAAAGTTACCGACATCAAAGCTGATATC
GCTACGAACAAAGATAATATTGCTAAAAAAGCAAACAGTGCCGACGTGTACACCAGAGAAGAGTCTGACAGCAAATTT
GTCAGAATTGATGGTCTGAACGCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCCATTGCCGAT
CACGATACTCGCCTGAACGGTTTGGATAAAACAGTGTCAGACCTGCGCAAAGAAACCCGCCAAGGCCTTGCAGAACAA
GCCGCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTCGGTTCAATGTAACGGCTGCAGTCGGCGGCTACAAATCC
GAATCGGCAGTCGCCATCGGTACCGGCTTCCGCTTTACCGAAAACTTTGCCGCCAAAGCAGGCGTGGCAGTCGGCACT
TCGTCCGGTTCTTCCGCAGCCTACCATGTCGGCGTCAATTACGAGTGGCTCGAGCACCACCACCACCACCACTGA
```

```
  1   MVAADIGAGL ADALTAPLDH KDKGLQSLTL DQSVRKNEKL KLAAQGAEKT
 51   YGNGDSLNTG KLKNDKVSRF DFIRQIEVDG QLITLESGEF QVYKQSHSAL
101   TAFQTEQIQD SEHSGKMVAK RQFRIGDIAG EHTSFDKLPE GGRATYRGTA
151   FGSDDAGGKL TYTIDFAAKQ GNGKIEHLKS PELNVDLAAA DIKPDGKRHA
201   VISGSVLYNQ AEKGSYSLGI FGGKAQEVAG SAEVKTVNGI RHIGLAAKQL
251   EGGGGTGSAT NDDDVKKAAT VAIAAAYNNG QEINGFKAGE TIYDIDEDGT
301   ITKKDATAAD VEADDFKGLG LKKVVTNLTK TVNENKQNVD AKVKAAESEI
351   EKLTTKLADT DAALADTDAA LDATTNALNK LGENITTFAE ETKTNIVKID
401   EKLEAVADTV DKHAEAFNDI ADSLDETNTK ADEAVKTANE AKQTAEETKQ
```

```
451     NVDAKVKAAE TAAGKAEAAA GTANTAADKA EAVAAKVTDI KADIATNKDN
501     IAKKANSADV YTREESDSKF VRIDGLNATT EKLDTRLASA EKSIADHDTR
551     LNGLDKTVSD LRKETRQGLA EQAALSGLFQ PYNVGRFNVT AAVGGYKSES
601     AVAIGTGFRF TENFAAKAGV AVGTSSGSSA AYHVGVNYEW LEHHHHHH*
```

ΔG741-961c

```
ATGGTCGCCGCCGACATCGGTGCGGGGCTTGCCGATGCACTAACCGCACCGCTCGACCATAAAGACAAAGGTTTGCAG
TCTTTGACGCTGGATCAGTCCGTCAGGAAAAACGAGAAACTGAAGCTGGCGGCACAAGGTGCGGAAAAAACTTATGGA
AACGGTGACAGCCTCAATACGGGCAAATTGAAGAACGACAAGGTCAGCCGTTTCGACTTTATCCGCCAAATCGAAGTG
GACGGGCAGCTCATTACCTTGGAGAGTGGAGAGTTCCAAGTATACAAACAAAGCCATTCCGCCTTAACCGCCTTTCAG
ACCGAGCAAATACAAGATTCGGAGCATTCCGGGAAGATGGTTGCGAAACGCCAGTTCAGAATCGGCGACATAGCGGGC
GAACATACATCTTTTGACAAGCTTCCCGAAGGCGGCAGGGCGACATATCGCGGGACGGCGTTCGGTTCAGACGATGCC
GGCGGAAAACTGACCTACACCATAGATTTCGCCGCCAAGCAGGGAAACGGCAAAATCGAACATTTGAAATCGCCAGAA
CTCAATGTCGACCTGGCCGCCGCCGATATCAAGCCGGATGGAAAACGCCATGCCGTCATCAGCGGTTCCGTCCTTTAC
AACCAAGCCGAGAAAGGCAGTTACTCCCTCGGTATCTTTGGCGGAAAAGCCCAGGAAGTTGCCGGCAGCGCGGAAGTG
AAAACCGTAAACGGCATACGCCATATCGGCCTTGCCGCCAAGCAACTCGAGGGTGGCGGAGGCACTGGATCCGCCACA
AACGACGACGATGTTAAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATCAACGGTTTC
AAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCCGATGTTGAA
GCCGACGACTTTAAAGGTCTGGGTCTGAAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAACAAACAAAAC
GTCGATGCCAAAGTAAAAGCTGCAGAATCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGATGCCGCTTTA
GCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAATTGGGAGAAAATATAACGACATTTGCTGAA
GAGACTAAGACAAATATCGTAAAAATTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCATGCCGAAGCA
TTCAACGATATCGCCGATTCATTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAATGAAGCCAAA
CAGACGGCCGAAGAAACCAAACAAAACGTCGATGCCAAAGTAAAAGCTGCAGAAACTGCAGCAGGCAAAGCCGAAGCT
GCCGCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTCGCTGCAAAAGTTACCGACATCAAAGCTGATATC
GCTACGAACAAAGATAATATTGCTAAAAAAGCAAACAGTGCCGACGTGTACACCAGAGAAGAGTCTGACAGCAAATTT
GTCAGAATTGATGGTCTGAACGCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCCATTGCCGAT
CACGATACTCGCCTGAACGGTTTGGATAAAACAGTGTCAGACCTGCGCAAAGAAACCCGCCAAGGCCTTGCAGAACAA
GCCGCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTCTCGAGCACCACCACCACCACCACTGA
```

```
  1     MVAADIGAGL ADALTAPLDH KDKGLQSLTL DQSVRKNEKL KLAAQGAEKT
 51     YGNGDSLNTG KLKNDKVSRF DFIRQIEVDG QLITLESGEF QVYKQSHSAL
101     TAFQTEQIQD SEHSGKMVAK RQFRIGDIAG EHTSFDKLPE GGRATYRGTA
151     FGSDDAGGKL TYTIDFAAKQ GNGKIEHLKS PELNVDLAAA DIKPDGKRHA
201     VISGSVLYNQ AEKGSYSLGI FGGKAQEVAG SAEVKTVNGI RHIGLAAKQL
251     EGGGGTGSAT NDDDVKKAAT VAIAAAYNNG QEINGFKAGE TIYDIDEDGT
301     ITKKDATAAD VEADDFKGLG LKKVVTNLTK TVNENKQNVD AKVKAAESEI
351     EKLTTKLADT DAALADTDAA LDATTNALNK LGENITTFAE ETKTNIVKID
401     EKLEAVADTV DKHAEAFNDI ADSLDETNTK ADEAVKTANE AKQTAEETKQ
451     NVDAKVKAAE TAAGKAEAAA GTANTAADKA EAVAAKVTDI KADIATNKDN
501     IAKKANSADV YTREESDSKF VRIDGLNATT EKLDTRLASA EKSIADHDTR
551     LNGLDKTVSD LRKETRQGLA EQAALSGLFQ PYNVGLEHHH HHH*
```

ΔG741-983

```
ATGGTCGCCGCCGACATCGGTGCGGGGCTTGCCGATGCACTAACCGCACCGCTCGACCATAAAGACAAAGGTTTGCAG
TCTTTGACGCTGGATCAGTCCGTCAGGAAAAACGAGAAACTGAAGCTGGCGGCACAAGGTGCGGAAAAAACTTATGGA
AACGGTGACAGCCTCAATACGGGCAAATTGAAGAACGACAAGGTCAGCCGTTTCGACTTTATCCGCCAAATCGAAGTG
GACGGGCAGCTCATTACCTTGGAGAGTGGAGAGTTCCAAGTATACAAACAAAGCCATTCCGCCTTAACCGCCTTTCAG
ACCGAGCAAATACAAGATTCGGAGCATTCCGGGAAGATGGTTGCGAAACGCCAGTTCAGAATCGGCGACATAGCGGGC
GAACATACATCTTTTGACAAGCTTCCCGAAGGCGGCAGGGCGACATATCGCGGGACGGCGTTCGGTTCAGACGATGCC
GGCGGAAAACTGACCTACACCATAGATTTCGCCGCCAAGCAGGGAAACGGCAAAATCGAACATTTGAAATCGCCAGAA
CTCAATGTCGACCTGGCCGCCGCCGATATCAAGCCGGATGGAAAACGCCATGCCGTCATCAGCGGTTCCGTCCTTTAC
AACCAAGCCGAGAAAGGCAGTTACTCCCTCGGTATCTTTGGCGGAAAAGCCCAGGAAGTTGCCGGCAGCGCGGAAGTG
AAAACCGTAAACGGCATACGCCATATCGGCCTTGCCGCCAAGCAACTCGAGGGATCCGGCGGAGGCGGCACTTCTGCG
CCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTAC
GCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGG
GATGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAATTTG
ATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGACACAGGCGAATCC
GTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCG
TATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAG
ACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCCCATATTATTGGCGGG
CGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACGAATGATGAAACC
AAGAACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAAC
AGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCG
TTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACCTGTCC
```

TACCACATCCGTAATAAAAACATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCC
CTATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTC
AAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCATTGCGGAATTACTGCCATGTGG
TGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCC
GCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACG
TTGCTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAG
GCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCC
TTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACC
TATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGCGTCGAAACC
AAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACC
GACCAATCCGGCGCAAACGAAACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGT
TTGGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAGGGGGCA
GGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGGCAGGATTATTCTTTCTTCACA
AACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTG
TCCTATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACAC
GCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAG
ACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCA
GCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACC
GCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCACGGGTCTGCGC
GTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAA
ACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGC
GAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCTAT
CTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGC
AGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACG
GTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGC
AACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTG
TTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCA
GCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATTCGGC
AACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGAGTCGGCGTA
GGCTACCGGTTCCTCGAGCACCACCACCACCACCACTGA

```
   1  MVAADIGAGL ADALTAPLDH KDKGLQSLTL DQSVRKNEKL KLAAQGAEKT
  51  YGNGDSLNTG KLKNDKVSRF DFIRQIEVDG QLITLESGEF QVYKQSHSAL
 101  TAFQTEQIQD SEHSGKMVAK RQFRIGDIAG EHTSFDKLPE GGRATYRGTA
 151  FGSDDAGGKL TYTIDFAAKQ GNGKIEHLKS PELNVDLAAA DIKPDGKRHA
 201  VISGSVLYNQ AEKGSYSLGI FGGKAQEVAG SAEVKTVNGI RHIGLAAKQL
 251  EGSGGGGTSA PDFNAGGTGI GSNSRATTAK SAAVSYAGIK NEMCKDRSML
 301  CAGRDDVAVT DRDAKINAPP PNLHTGDFPN PNDAYKNLIN LKPAIEAGYT
 351  GRGVEVGIVD TGESVGSISF PELYGRKEHG YNENYKNYTA YMRKEAPEDG
 401  GGKDIEASFD DEAVIETEAK PTDIRHVKEI GHIDLVSHII GGRSVDGRPA
 451  GGIAPDATLH IMNTNDETKN EMMVAAIRNA WVKLGERGVR IVNNSFGTTS
 501  RAGTADLFQI ANSEEQYRQA LLDYSGGDKT DEGIRLMQQS DYGNLSYHIR
 551  NKNMLFIFST GNDAQAQPNT YALLPFYEKD AQKGIITVAG VDRSGEKFKR
 601  EMYGEPGTEP LEYGSNHCGI TAMWCLSAPY EASVRFTRTN PIQIAGTSFS
 651  APIVTGTAAL LLQKYPWMSN DNLRTTLLTT AQDIGAVGVD SKFGWGLLDA
 701  GKAMNGPASF PFGDFTADTK GTSDIAYSFR NDISGTGGLI KKGGSQLQLH
 751  GNNTYTGKTI IEGGSLVLYG NNKSDMRVET KGALIYNGAA SGGSLNSDGI
 801  VYLADTDQSG ANETVHIKGS LQLDGKGTLY TRLGKLLKVD GTAIIGGKLY
 851  MSARGKGAGY LNSTGRRVPF LSAAKIGQDY SFFTNIETDG GLLASLDSVE
 901  KTAGSEGDTL SYYVRRGNAA RTASAAAHSA PAGLKHAVEQ GGSNLENLMV
 951  ELDASESSAT PETVETAAAD RTDMPGIRPY GATFRAAAAV QHANAADGVR
1001  IFNSLAATVY ADSTAAHADM QGRRLKAVSD GLDHNGTGLR VIAQTQQDGG
1051  TWEQGGVEGK MRGSTQTVGI AAKTGENTTA AATLGMGRST WSENSANAKT
1101  DSISLFAGIR HDAGDIGYLK GLFSYGRYKN SISRSTGADE HAEGSVNGTL
1151  MQLGALGGVN VPFAATGDLT VEGGLRYDLL KQDAFAEKGS ALGWSGNSLT
1201  EGTLVGLAGL KLSQPLSDKA VLFATAGVER DLNGRDYTVT GGFTGATAAT
1251  GKTGARNMPH TRLVAGLGAD VEFGNGWNGL ARYSYAGSKQ YGNHSGRVGV
1301  GYRFLEHHHH HH*
```

ΔG741–ORF46.1

ATGGTCGCCGCCGACATCGGTGCGGGGCTTGCCGATGCACTAACCGCACCGCTCGACCATAAAGACAAAGGTTTGCAG
TCTTTGACGCTGGATCAGTCCGTCAGGAAAAACGAGAAACTGAAGCTGGCGGCACAAGGTGCGGAAAAAACTTATGGA
AACGGTGACAGCCTCAATACGGGCAAATTGAAGAACGACAAGGTCAGCCGTTTCGACTTTATCCGCCAAATCGAAGTG
GACGGGCAGCTCATTACCTTGGAGAGTGGAGAGTTCCAAGTATACAAACAAAGCCATTCCGCCTTAACCGCCTTTCAG
ACCGAGCAAATACAAGATTCGGAGCATTCCGGGAAGATGGTTGCGAAACGCCAGTTCAGAATCGGCGACATAGCGGGC
GAACATACATCTTTTGACAAGCTTCCCGAAGGCGGCAGGGCGACATATCGCGGGACGGCGTTCGGTTCAGACGATGCC

```
GGCGGAAAACTGACCTACACCATAGATTTCGCCGCCAAGCAGGGAAACGGCAAAATCGAACATTTGAAATCGCCAGAA
CTCAATGTCGACCTGGCCGCCGCCGATATCAAGCCGGATGGAAAACGCCATGCCGTCATCAGCGGTTCCGTCCTTTAC
AACCAAGCCGAGAAAGGCAGTTACTCCCTCGGTATCTTTGGCGGAAAAGCCCAGGAAGTTGCCGGCAGCGCGGAAGTG
AAAACCGTAAACGGCATACGCCATATCGGCCTTGCCGCCAAGCAACTCGACGGTGGCGGAGGCACTGGATCCTCAGAT
TTGGCAAACGATTCTTTTATCCGGCAGGTTCTCGACCGTCAGCATTTCGAACCCGACGGAAATACCACCTATTCGGC
AGCAGGGGGGAACTTGCCGAGCGCAGCGGCCATATCGGATTGGGAAAAATACAAAGCCATCAGTTGGGCAACCTGATG
ATTCAACAGGCGGCCATTAAAGGAAATATCGGCTACATTGTCCGCTTTTCCGATCACGGGCACGAAGTCCATTCCCCC
TTCGACAACCATGCCTCACATTCCGATTCTGATGAAGCCGGTAGTCCCGTTGACGGATTTAGCCTTTACCGCATCCAT
TGGGACGGATACGAACACCATCCCGCCGACGGCTATGACGGGCCACAGGGCGGCGGCTATCCCGCTCCCAAAGGCGCG
AGGGATATATACAGCTACGACATAAAAGGCGTTGCCCAAAATATCCGCCTCAACCTGACCGACAACCGCAGCACCGGA
CAACGGCTTGCCGACCGTTTCCACAATGCCGGTAGTATGCTGACGCAAGGAGTAGGCGACGGATTCAAACGCGCCACC
CGATACAGCCCCGAGCTGGACAGATCGGGCAATGCCGCCGAAGCCTTCAACGGCACTGCAGATATCGTTAAAAACATC
ATCGGCGCGGCAGGAGAAATTGTCGGCGCAGGCGATGCCGTGCAGGGCATAAGCGAAGGCTCAAACATTGCTGTCATG
CACGGCTTGGGTCTGCTTTCCACCGAAAACAAGATGGCGCGCATCAACGATTTGGCAGATATGGCGCAACTCAAAGAC
TATGCCGCAGCAGCCATCCGCGATTGGGCAGTCCAAAACCCCAATGCCGCACAAGGCATAGAAGCCGTCAGCAATATC
TTTATGGCAGCCATCCCCATCAAAGGGATTGGAGCTGTTCGGGGAAAATACGGCTTGGGCGGCATCACGGCACATCCT
ATCAAGCGGTCGCAGATGGGCGCGATCGCATTGCCGAAAGGGAAATCCGCCGTCAGCGACAATTTTGCCGATGCGGCA
TACGCCAAATACCCGTCCCCTTACCATTCCCGAAATATCCGTTCAAACTTGGAGCAGCGTTACGGCAAAGAAAACATC
ACCTCCTCAACCGTGCCGCCGTCAAACGGCAAAAATGTCAAACTGGCAGACCAACGCCACCCGAAGACAGGCGTACCG
TTTGACGGTAAAGGGTTTCCGAATTTTGAGAAGCACGTGAAATATGATACGCTCGAGCACCACCACCACCACCACTGA
```

```
  1  MVAADIGAGL ADALTAPLDH KDKGLQSLTL DQSVRKNEKL KLAAQGAEKT
 51  YGNGDSLNTG KLKNDKVSRF DFIRQIEVDG QLITLESGEF QVYKQSHSAL
101  TAFQTEQIQD SEHSGKMVAK RQFRIGDIAG EHTSFDKLPE GGRATYRGTA
151  FGSDDAGGKL TYTIDFAAKQ GNGKIEHLKS PELNVDLAAA DIKPDGKRHA
201  VISGSVLYNQ AEKGSYSLGI FGGKAQEVAG SAEVKTVNGI RHIGLAAKQL
251  DGGGGTGSSD LANDSFIRQV LDRQHFEPDG KYHLFGSRGE LAERSGHIGL
301  GKIQSHQLGN LMIQQAAIKG NIGYIVRFSD HGHEVHSPFD NHASHSDSDE
351  AGSPVDGFSL YRIHWDGYEH HPADGYDGPQ GGGYPAPKGA RDIYSYDIKG
401  VAQNIRLNLT DNRSTGQRLA DRFHNAGSML TQGVGDGFKR ATRYSPELDR
451  SGNAAEAFNG TADIVKNIIG AAGEIVGAGD AVQGISEGSN IAVMHGLGLL
501  STENKMARIN DLADMAQLKD YAAAAIRDWA VQNPNAAQGI EAVSNIFMAA
551  IPIKGIGAVR GKYGLGGITA HPIKRSQMGA IALPKGKSAV SDNFADAAYA
601  KYPSPYHSRN IRSNLEQRYG KENITSSTVP PSNGKNVKLA DQRHPKTGVP
651  FDGKGFPNFE KHVKYDTLEH HHHHH*
```

Example 5 – hybrids of 287

Expression of 287 as full-length with a C-terminal His-tag, or without its leader peptide but with a C-terminal His-tag, gives fairly low expression levels. Better expression is achieved using a N-terminal GST-fusion. As an alternative to using GST as an N-terminal fusion partner, 287 was placed at the C-terminus of protein 919 ('919-287'), of protein 953 ('953-287'), and of proteins ORF46.1 ('ORF46.1-287'). In both cases, the leader peptides were deleted, and the hybrids were direct in-frame fusions.

To generate the 953-287 hybrid, the leader peptides of the two proteins were omitted by designing the forward primer downstream from the leader of each sequence; the stop codon sequence was omitted in the 953 reverse primer but included in the 287 reverse primer. For the 953 gene, the 5' and the 3' primers used for amplification included a *Nde*I and a *Bam*HI restriction sites respectively, whereas for the amplification of the 287 gene the 5' and the 3' primers included a *Bam*HI and a *Xho*I restriction sites respectively. In this way a sequential directional cloning of the two genes in pET21b+, using *Nde*I-*Bam*HI (to clone the first gene) and subsequently *Bam*HI-*Xho*I (to clone the second gene) could be achieved.

The 919-287 hybrid was obtained by cloning the sequence coding for the mature portion of 287 into the *Xho*I site at the 3'-end of the 919-His clone in pET21b+. The primers used for amplification of the 287 gene were designed for introducing a *Sal*I restriction site at the 5'- and a *Xho*I site at the 3'- of the PCR fragment. Since the cohesive ends produced by the *Sal*I and *Xho*I restriction enzymes are compatible, the 287 PCR product digested with *Sal*I-*Xho*I could be inserted in the pET21b-919 clone cleaved with *Xho*I.

The ORF46.1-287 hybrid was obtained similarly.

The bactericidal efficacy (homologous strain) of antibodies raised against the hybrid proteins was compared with antibodies raised against simple mixtures of the component antigens:

	Mixture with 287	Hybrid with 287
919	32000	16000
953	8192	8192
ORF46.1	128	8192

Data for bactericidal activity against heterologous MenB strains and against serotypes A and C were also obtained for 919-287 and 953-287:

	919		953		ORF46.1	
Strain	*Mixture*	*Hybrid*	*Mixture*	*Hybrid*	*Mixture*	*Hybrid*
MC58	512	1024	512	1024	-	1024
NGH38	1024	2048	2048	4096	-	4096
BZ232	512	128	1024	16	-	-
MenA (F6124)	512	2048	2048	32	-	1024
MenC (C11)	>2048	n.d.	>2048	n.d.	-	n.d.
MenC (BZ133)	>4096	>8192	>4096	<16	-	2048

Hybrids of ORF46.1 and 919 were also constructed. Best results (four-fold higher titre) were achieved with 919 at the N-terminus.

Hybrids 919-519His, ORF97-225His and 225-ORF97His were also tested. These gave moderate ELISA titres and bactericidal antibody responses.

As hybrids of two proteins A & B may be either NH_2–A–B–COOH or NH_2–B–A–COOH, the "reverse" hybrids with 287 at the N-terminus were also made, but using ΔG287. A panel of strains was used, including homologous strain 2996. FCA was used as adjuvant:

	287 & 919		287 & 953		287 & ORF46.1	
Strain	*ΔG287-919*	*919-287*	*ΔG287-953*	*953-287*	*ΔG287-46.1*	*46.1-287*
2996	128000	16000	65536	8192	16384	8192
BZ232	256	128	128	<4	<4	<4
1000	2048	<4	<4	<4	<4	<4
MC58	8192	1024	16384	1024	512	128
NGH38	32000	2048	>2048	4096	16384	4096
394/98	4096	32	256	128	128	16
MenA (F6124)	32000	2048	>2048	32	8192	1024
MenC (BZ133)	64000	>8192	>8192	<16	8192	2048

Better bactericidal titres are generally seen with 287 at the N-terminus.

When fused to protein 961 [NH_2-ΔG287-961-COOH – sequence shown above], the resulting protein is insoluble and must be denatured and renatured for purification. Following renaturation, around 50% of the protein was found to remain insoluble. The soluble and insoluble proteins were compared, and much better bactericidal titres were obtained with the soluble protein (FCA as adjuvant):

	2996	BZ232	MC58	NGH38	F6124	BZ133
Soluble	65536	128	4096	>2048	>2048	4096
Insoluble	8192	<4	<4	16	n.d.	n.d.

Titres with the insoluble form were, however, improved by using alum adjuvant instead:

Insoluble	32768	128	4096	>2048	>2048	2048

961c was also used in hybrid proteins (see above). As 961 and its domain variants direct efficient expression, they are ideally suited as the N-terminal portion of a hybrid protein.

Example 23 – further hybrids

Further hybrid proteins of the invention are shown in the drawings and have the sequences set out below. These are advantageous when compared to the individual proteins:

<u>ORF46.1-741</u>

```
ATGTCAGATTTGGCAAACGATTCTTTTATCCGGCAGGTTCTCGACCGTCAGCATTTCGAACCCGACGGAAATACCAC
CTATTCGGCAGCAGGGGGAACTTGCCGAGCGCAGCGGCCATATCGGATTGGGAAAAATACAAAGCCATCAGTTGGGC
AACCTGATGATTCAACAGGCGGCCATTAAAGGAAATATCGGCTACATTGTCCGCTTTTCCGATCACGGGCACGAAGTC
CATTCCCCCTTCGACAACCATGCCTCACATTCCGATTCTGATGAAGCCGGTAGTCCCGTTGACGGATTTAGCCTTTAC
CGCATCCATTGGGACGGATACGAACACCATCCCGCCGACGGCTATGACGGGCCACAGGGCGGCGGCTATCCCGCTCCC
AAAGGCGCGAGGGATATATACAGCTACGACATAAAAGGCGTTGCCCAAAATATCCGCCTCAACCTGACCGACAACCGC
AGCACCGGACAACGGCTTGCCGACCGTTTCCACAATGCCGGTAGTATGCTGACGCAAGGAGTAGGCGACGGATTCAAA
CGCGCCACCCGATACAGCCCCGAGCTGGACAGATCGGGCAATGCCGCCGAAGCCTTCAACGGCACTGCAGATATCGTT
AAAAACATCATCGGCGCGGCAGGAGAAATTGTCGGCGCAGGCGATGCCGTGCAGGGCATAAGCGAAGGCTCAAACATT
```

GCTGTCATGCACGGCTTGGGTCTGCTTTCCACCGAAAACAAGATGGCGCGCATCAACGATTTGGCAGATATGGCGCAA
CTCAAAGACTATGCCGCAGCAGCCATCCGCGATTGGGCAGTCCAAAACCCCAATGCCGCACAAGGCATAGAAGCCGTC
AGCAATATCTTTATGGCAGCCATCCCCATCAAAGGGATTGGAGCTGTTCGGGGAAAATACGGCTTGGGCGGCATCACG
GCACATCCTATCAAGCGGTCGCAGATGGGCGCGATCGCATTGCCGAAAGGGAAATCCGCCGTCAGCGACAATTTTGCC
GATGCGGCATACGCCAAATACCCGTCCCCTTACCATTCCCGAAATATCCGTTCAAACTTGGAGCAGCGTTACGGCAAA
GAAAACATCACCTCCTCAACCGTGCCGCCGTCAAACGGCAAAAATGTCAAACTGGCAGACCAACGCCACCCGAAGACA
GGCGTACCGTTTGACGGTAAAGGGTTTCCGAATTTTGAGAAGCACGTGAAATATGATACGGGATCCGGAGGGGGTGGT
GTCGCCGCCGACATCGGTGCGGGGCTTGCCGATGCACTAACCGCACCGCTCGACCATAAAGACAAAGGTTTGCAGTCT
TTGACGCTGGATCAGTCCGTCAGGAAAAACGAGAAACTGAAGCTGGCGGCACAAGGTGCGGAAAAAACTTATGGAAAC
GGTGACAGCCTCAATACGGGCAAATTGAAGAACGACAAGGTCAGCCGTTTCGACTTTATCCGCCAAATCGAAGTGGAC
GGGCAGCTCATTACCTTGGAGAGTGGAGAGTTCCAAGTATACAAACAAAGCCATTCCGCCTTAACCGCCTTTCAGACC
GAGCAAATACAAGATTCGGAGCATTCCGGGAAGATGGTTGCGAAACGCCAGTTCAGAATCGGCGACATAGCGGGCGAA
CATACATCTTTTGACAAGCTTCCCGAAGGCGGCAGGGCGACATATCGCGGGACGGCGTTCGGTTCAGACGATGCCGGC
GGAAAACTGACCTACACCATAGATTTCGCCGCCAAGCAGGGAAACGGCAAAATCGAACATTTGAAATCGCCAGAACTC
AATGTCGACCTGGCCGCCGCCGATATCAAGCCGGATGGAAAACGCCATGCCGTCATCAGCGGTTCCGTCCTTTACAAC
CAAGCCGAGAAAGGCAGTTACTCCCTCGGTATCTTTGGCGGAAAAGCCCAGGAAGTTGCCGGCAGCGCGGAAGTGAAA
ACCGTAAACGGCATACGCCATATCGGCCTTGCCGCCAAGCAACTCGAGCACCACCACCACCACCACTGA

```
  1  MSDLANDSFI RQVLDRQHFE PDGKYHLFGS RGELAERSGH IGLGKIQSHQ
 51  LGNLMIQQAA IKGNIGYIVR FSDHGHEVHS PFDNHASHSD SDEAGSPVDG
101  FSLYRIHWDG YEHHPADGYD GPQGGGYPAP KGARDIYSYD IKGVAQNIRL
151  NLTDNRSTGQ RLADRFHNAG SMLTQGVGDG FKRATRYSPE LDRSGNAAEA
201  FNGTADIVKN IIGAAGEIVG AGDAVQGISE GSNIAVMHGL GLLSTENKMA
251  RINDLADMAQ LKDYAAAAIR DWAVQNPNAA QGIEAVSNIF MAAIPIKGIG
301  AVRGKYGLGG ITAHPIKRSQ MGAIALPKGK SAVSDNFADA AYAKYPSPYH
351  SRNIRSNLEQ RYGKENITSS TVPPSNGKNV KLADQRHPKT GVPFDGKGFP
401  NFEKHVKYDT GSGGGVAAD  IGAGLADALT APLDHKDKGL QSLTLDQSVR
451  KNEKLKLAAQ GAEKTYGNGD SLNTGKLKND KVSRFDFIRQ IEVDGQLITL
501  ESGEFQVYKQ SHSALTAFQT EQIQDSEHSG KMVAKRQFRI GDIAGEHTSF
551  DKLPEGGRAT YRGTAFGSDD AGGKLTYTID FAAKQGNGKI EHLKSPELNV
601  DLAAADIKPD GKRHAVISGS VLYNQAEKGS YSLGIFGGKA QEVAGSAEVK
651  TVNGIRHIGL AAKQLEHHHH HH*
```

ORF46.1-961

ATGTCAGATTTGGCAAACGATTCTTTTATCCGGCAGGTTCTCGACCGTCAGCATTTCGAACCCGACGGGAAATACCAC
CTATTCGGCAGCAGGGGGGAACTTGCCGAGCGCAGCGGCCATATCGGATTGGGAAAAATACAAAGCCATCAGTTGGGC
AACCTGATGATTCAACAGGCGGCCATTAAAGGAAATATCGGCTACATTGTCCGCTTTTCCGATCACGGGCACGAAGTC
CATTCCCCCTTCGACAACCATGCCTCACATTCCGATTCTGATGAAGCCGGTAGTCCCGTTGACGGATTTAGCCTTTAC
CGCATCCATTGGGACGGATACGAACACCATCCCGCCGACGGCTATGACGGGCCACAGGGCGGCGGCTATCCCGCTCCC
AAAGGCGCGAGGGATATATACAGCTACGACATAAAAGGCGTTGCCCAAAATATCCGCCTCAACCTGACCGACAACCGC
AGCACCGGACAACGGCTTGCCGACCGTTTCCACAATGCCGGTAGTATGCTGACGCAAGGAGTAGGCGACGGATTCAAA
CGCGCCACCCGATACAGCCCCGAGCTGGACAGATCGGGCAATGCCGCCGAAGCCTTCAACGGCACTGCAGATATCGTT
AAAAACATCATCGGCGCGGCAGGAGAAATTGTCGGCGCAGGCGATGCCGTGCAGGGCATAAGCGAAGGCTCAAACATT
GCTGTCATGCACGGCTTGGGTCTGCTTTCCACCGAAAACAAGATGGCGCGCATCAACGATTTGGCAGATATGGCGCAA
CTCAAAGACTATGCCGCAGCAGCCATCCGCGATTGGGCAGTCCAAAACCCCAATGCCGCACAAGGCATAGAAGCCGTC
AGCAATATCTTTATGGCAGCCATCCCCATCAAAGGGATTGGAGCTGTTCGGGGAAAATACGGCTTGGGCGGCATCACG
GCACATCCTATCAAGCGGTCGCAGATGGGCGCGATCGCATTGCCGAAAGGGAAATCCGCCGTCAGCGACAATTTTGCC
GATGCGGCATACGCCAAATACCCGTCCCCTTACCATTCCCGAAATATCCGTTCAAACTTGGAGCAGCGTTACGGCAAA
GAAAACATCACCTCCTCAACCGTGCCGCCGTCAAACGGCAAAAATGTCAAACTGGCAGACCAACGCCACCCGAAGACA
GGCGTACCGTTTGACGGTAAAGGGTTTCCGAATTTTGAGAAGCACGTGAAATATGATACGGGATCCGGAGGAGGAGGA
GCCACAAACGACGACGATGTTAAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATCAAC
GGTTTCAAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCCGAT
GTTGAAGCCGACGACTTTAAAGGTCTGGGTCTGAAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAACAAA
CAAAACGTCGATGCCAAAGTAAAAGCTGCAGAATCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGATGCC
GCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAATTGGGAGAAAATATAACGACATTT
GCTGAAGAGACTAAGACAAATATCGTAAAAATTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCATGCC
GAAGCATTCAACGATATCGCCGATTCATTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAATGAA
GCCAAACAGACGGCCGAAGAAACCAAACAAAACGTCGATGCCAAAGTAAAAGCTGCAGAAACTGCAGCAGGCAAAGCC
GAAGCTGCCGCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTCGCTGCAAAAGTTACCGACATCAAAGCT
GATATCGCTACGAACAAAGATAATATTGCTAAAAAAGCAAACAGTGCCGACGTGTACACCAGAGAAGAGTCTGACAGC
AAATTTGTCAGAATTGATGGTCTGAACGCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCCATT
GCCGATCACGATACTCGCCTGAACGGTTTGGATAAAACAGTGTCAGACCTGCGCAAAGAAACCCGCCAAGGCCTTGCA
GAACAAGCCGCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTCGGTTCAATGTAACGGCTGCAGTCGGCGGCTAC
AAATCCGAATCGGCAGTCGCCATCGGTACCGGCTTCCGCTTTACCGAAAACTTTGCCGCCAAAGCAGGCGTGGCAGTC
GGCACTTCGTCCGGTTCTTCCGCAGCCTACCATGTCGGCGTCAATTACGAGTGGCTCGAGCACCACCACCACCACCAC
TGA

```
  1   MSDLANDSFI RQVLDRQHFE PDGKYHLFGS RGELAERSGH IGLGKIQSHQ
 51   LGNLMIQQAA IKGNIGYIVR FSDHGHEVHS PFDNHASHSD SDEAGSPVDG
101   FSLYRIHWDG YEHHPADGYD GPQGGGYPAP KGARDIYSYD IKGVAQNIRL
151   NLTDNRSTGQ RLADRFHNAG SMLTQGVGDG FKRATRYSPE LDRSGNAAEA
201   FNGTADIVKN IIGAAGEIVG AGDAVQGISE GSNIAVMHGL GLLSTENKMA
251   RINDLADMAQ LKDYAAAAIR DWAVQNPNAA QGIEAVSNIF MAAIPIKGIG
301   AVRGKYGLGG ITAHPIKRSQ MGAIALPKGK SAVSDNFADA AYAKYPSPYH
351   SRNIRSNLEQ RYGKENITSS TVPPSNGKNV KLADQRHPKT GVPFDGKGFP
401   NFEKHVKYDT GSGGGGATND DDVKKAATVA IAAAYNNGQE INGFKAGETI
451   YDIDEDGTIT KKDATAADVE ADDFKGLGLK KVVTNLTKTV NENKQNVDAK
501   VKAAESEIEK LTTKLADTDA ALADTDAALD ATTNALNKLG ENITTFAEET
551   KTNIVKIDEK LEAVADTVDK HAEAFNDIAD SLDETNTKAD EAVKTANEAK
601   QTAEETKQNV DAKVKAAETA AGKAEAAAGT ANTAADKAEA VAAKVTDIKA
651   DIATNKDNIA KKANSADVYT REESDSKFVR IDGLNATTEK LDTRLASAEK
701   SIADHDTRLN GLDKTVSDLR KETRQGLAEQ AALSGLFQPY NVGRFNVTAA
751   VGGYKSESAV AIGTGFRFTE NFAAKAGVAV GTSSGSSAAY HVGVNYEWLE
801   HHHHHH*
```

ORF46.1-961c

```
ATGTCAGATTTGGCAAACGATTCTTTTATCCGGCAGGTTCTCGACCGTCAGCATTTCGAACCCGACGGGAAATACCAC
CTATTCGGCAGCAGGGGGGAACTTGCCGAGCGCAGCGGCCATATCGGATTGGGAAAAATACAAAGCCATCAGTTGGGC
AACCTGATGATTCAACAGGCGGCCATTAAAGGAAATATCGGCTACATTGTCCGCTTTTCCGATCACGGGCACGAAGTC
CATTCCCCCTTCGACAACCATGCCTCACATTCCGATTCTGATGAAGCCGGTAGTCCCGTTGACGGATTTAGCCTTTAC
CGCATCCATTGGGACGGATACGAACACCATCCCGCCGACGGCTATGACGGGCCACAGGGCGGCGGCTATCCCGCTCCC
AAAGGCGCGAGGGATATATACAGCTACGACATAAAAGGCGTTGCCCAAAATATCCGCCTCAACCTGACCGACAACCGC
AGCACCGGACAACGGCTTGCCGACCGTTTCCACAATGCCGGTAGTATGCTGACGCAAGGAGTAGGCGACGGATTCAAA
CGCGCCACCCGATACAGCCCCGAGCTGGACAGATCGGGCAATGCCGCCGAAGCCTTCAACGGCACTGCAGATATCGTT
AAAAACATCATCGGCGCGGCAGGAGAAATTGTCGGCGCAGGCGATGCCGTGCAGGGCATAAGCGAAGGCTCAAACATT
GCTGTCATGCACGGCTTGGGTCTGCTTTCCACCGAAAACAAGATGGCGCGCATCAACGATTTGGCAGATATGGCGCAA
CTCAAAGACTATGCCGCAGCAGCCATCCGCGATTGGGCAGTCCAAAACCCCAATGCCGCACAAGGCATAGAAGCCGTC
AGCAATATCTTTATGGCAGCCATCCCCATCAAAGGGATTGGAGCTGTTCGGGGAAAATACGGCTTGGGCGGCATCACG
GCACATCCTATCAAGCGGTCGCAGATGGGCGCGATCGCATTGCCGAAAGGGAAATCCGCCGTCAGCGACAATTTTGCC
GATGCGGCATACGCCAAATACCCGTCCCCTTACCATTCCCGAAATATCCGTTCAAACTTGGAGCAGCGTTACGGCAAA
GAAAACATCACCTCCTCAACCGTGCCGCCGTCAAACGGCAAAAATGTCAAACTGGCAGACCAACGCCACCCGAAGACA
GGCGTACCGTTTGACGGTAAAGGGTTTCCGAATTTTGAGAAGCACGTGAAATATGATACGGGATCCGGAGGAGGAGGA
GCCACAAACGACGACGATGTTAAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATCAAC
GGTTTCAAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCCGAT
GTTGAAGCCGACGACTTTAAAGGTCTGGGTCTGAAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAACAAA
CAAAACGTCGATGCCAAAGTAAAAGCTGCAGAATCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGATGCC
GCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAATTGGGAGAAAATATAACGACATTT
GCTGAAGAGACTAAGACAAATATCGTAAAAATTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCATGCC
GAAGCATTCAACGATATCGCCGATTCATTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAATGAA
GCCAAACAGACGGCCGAAGAAACCAAACAAAACGTCGATGCCAAAGTAAAAGCTGCAGAAACTGCAGCAGGCAAAGCC
GAAGCTGCCGCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTCGCTGCAAAAGTTACCGACATCAAAGCT
GATATCGCTACGAACAAAGATAATATTGCTAAAAAAGCAAACAGTGCCGACGTGTACACCAGAGAAGAGTCTGACAGC
AAATTTGTCAGAATTGATGGTCTGAACGCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCCATT
GCCGATCACGATACTCGCCTGAACGGTTTGGATAAAACAGTGTCAGACCTGCGCAAAGAAACCCGCCAAGGCCTTGCA
GAACAAGCCGCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTCTCGAGCACCACCACCACCACCACTGA
```

```
  1   MSDLANDSFI RQVLDRQHFE PDGKYHLFGS RGELAERSGH IGLGKIQSHQ
 51   LGNLMIQQAA IKGNIGYIVR FSDHGHEVHS PFDNHASHSD SDEAGSPVDG
101   FSLYRIHWDG YEHHPADGYD GPQGGGYPAP KGARDIYSYD IKGVAQNIRL
151   NLTDNRSTGQ RLADRFHNAG SMLTQGVGDG FKRATRYSPE LDRSGNAAEA
201   FNGTADIVKN IIGAAGEIVG AGDAVQGISE GSNIAVMHGL GLLSTENKMA
251   RINDLADMAQ LKDYAAAAIR DWAVQNPNAA QGIEAVSNIF MAAIPIKGIG
301   AVRGKYGLGG ITAHPIKRSQ MGAIALPKGK SAVSDNFADA AYAKYPSPYH
351   SRNIRSNLEQ RYGKENITSS TVPPSNGKNV KLADQRHPKT GVPFDGKGFP
401   NFEKHVKYDT GSGGGGATND DDVKKAATVA IAAAYNNGQE INGFKAGETI
451   YDIDEDGTIT KKDATAADVE ADDFKGLGLK KVVTNLTKTV NENKQNVDAK
501   VKAAESEIEK LTTKLADTDA ALADTDAALD ATTNALNKLG ENITTFAEET
551   KTNIVKIDEK LEAVADTVDK HAEAFNDIAD SLDETNTKAD EAVKTANEAK
601   QTAEETKQNV DAKVKAAETA AGKAEAAAGT ANTAADKAEA VAAKVTDIKA
651   DIATNKDNIA KKANSADVYT REESDSKFVR IDGLNATTEK LDTRLASAEK
701   SIADHDTRLN GLDKTVSDLR KETRQGLAEQ AALSGLFQPY NVGLEHHHHH
751   H*
```

961-ORF46.1

ATGGCCACAAACGACGACGATGTTAAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATC
AACGGTTTCAAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCC
GATGTTGAAGCCGACGACTTTAAAGGTCTGGGTCTGAAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAAC
AAACAAAACGTCGATGCCAAAGTAAAAGCTGCAGAATCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGAT
GCCGCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAATTGGGAGAAAATATAACGACA
TTTGCTGAAGAGACTAAGACAAATATCGTAAAAATTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCAT
GCCGAAGCATTCAACGATATCGCCGATTCATTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAAT
GAAGCCAAACAGACGGCCGAAGAAACCAAACAAAACGTCGATGCCAAAGTAAAAGCTGCAGAAACTGCAGCAGGCAAA
GCCGAAGCTGCCGCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTCGCTGCAAAAGTTACCGACATCAAA
GCTGATATCGCTACGAACAAAGATAATATTGCTAAAAAAGCAAACAGTGCCGACGTGTACACCAGAGAAGAGTCTGAC
AGCAAATTTGTCAGAATTGATGGTCTGAACGCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCC
ATTGCCGATCACGATACTCGCCTGAACGGTTTGGATAAAACAGTGTCAGACCTGCGCAAAGAAACCCGCCAAGGCCTT
GCAGAACAAGCCGCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTCGGTTCAATGTAACGGCTGCAGTCGGCGGC
TACAAATCCGAATCGGCAGTCGCCATCGGTACCGGCTTCCGCTTTACCGAAAACTTTGCCGCCAAAGCAGGCGTGGCA
GTCGGCACTTCGTCCGGTTCTTCCGCAGCCTACCATGTCGGCGTCAATTACGAGTGGGGATCCGGAGGAGGAGGATCA
GATTTGGCAAACGATTCTTTTATCCGGCAGGTTCTCGACCGTCAGCATTTCGAACCCGACGGGAAATACCACCTATTC
GGCAGCAGGGGGGAACTTGCCGAGCGCAGCGGCCATATCGGATTGGGAAAAATACAAAGCCATCAGTTGGGCAACCTG
ATGATTCAACAGGCGGCCATTAAAGGAAATATCGGCTACATTGTCCGCTTTTCCGATCACGGGCACGAAGTCCATTCC
CCCTTCGACAACCATGCCTCACATTCCGATTCTGATGAAGCCGGTAGTCCCGTTGACGGATTTAGCCTTTACCGCATC
CATTGGGACGGATACGAACACCATCCCGCCGACGGCTATGACGGGCCACAGGGCGGCGGCTATCCCGCTCCCAAAGGC
GCGAGGGATATATACAGCTACGACATAAAAGGCGTTGCCCAAAATATCCGCCTCAACCTGACCGACAACCGCAGCACC
GGACAACGGCTTGCCGACCGTTTCCACAATGCCGGTAGTATGCTGACGCAAGGAGTAGGCGACGGATTCAAACGCGCC
ACCCGATACAGCCCCGAGCTGGACAGATCGGGCAATGCCGCCGAAGCCTTCAACGGCACTGCAGATATCGTTAAAAAC
ATCATCGGCGCGGCAGGAGAAATTGTCGGCGCAGGCGATGCCGTGCAGGGCATAAGCGAAGGCTCAAACATTGCTGTC
ATGCACGGCTTGGGTCTGCTTTCCACCGAAAACAAGATGGCGCGCATCAACGATTTGGCAGATATGGCGCAACTCAAA
GACTATGCCGCAGCAGCCATCCGCGATTGGGCAGTCCAAAACCCCAATGCCGCACAAGGCATAGAAGCCGTCAGCAAT
ATCTTTATGGCAGCCATCCCCATCAAAGGGATTGGAGCTGTTCGGGGAAAATACGGCTTGGGCGGCATCACGGCACAT
CCTATCAAGCGGTCGCAGATGGGCGCGATCGCATTGCCGAAAGGGAAATCCGCCGTCAGCGACAATTTTGCCGATGCG
GCATACGCCAAATACCCGTCCCCTTACCATTCCCGAAATATCCGTTCAAACTTGGAGCAGCGTTACGGCAAAGAAAAC
ATCACCTCCTCAACCGTGCCGCCGTCAAACGGCAAAAATGTCAAACTGGCAGACCAACGCCACCCGAAGACAGGCGTA
CCGTTTGACGGTAAAGGGTTTCCGAATTTTGAGAAGCACGTGAAATATGATACGCTCGAGCACCACCACCACCACCAC
TGA

1	MATNDDDVKK	AATVAIAAAY	NNGQEINGFK	AGETIYDIDE	DGTITKKDAT
51	AADVEADDFK	GLGLKKVVTN	LTKTVNENKQ	NVDAKVKAAE	SEIEKLTTKL
101	ADTDAALADT	DAALDATTNA	LNKLGENITT	FAEETKTNIV	KIDEKLEAVA
151	DTVDKHAEAF	NDIADSLDET	NTKADEAVKT	ANEAKQTAEE	TKQNVDAKVK
201	AAETAAGKAE	AAAGTANTAA	DKAEAVAAKV	TDIKADIATN	KDNIAKKANS
251	ADVYTREESD	SKFVRIDGLN	ATTEKLDTRL	ASAEKSIADH	DTRLNGLDKT
301	VSDLRKETRQ	GLAEQAALSG	LFQPYNVGRF	NVTAAVGGYK	SESAVAIGTG
351	FRFTENFAAK	AGVAVGTSSG	SSAAYHVGVN	YEWGSGGGGS	DLANDSFIRQ
401	VLDRQHFEPD	GKYHLFGSRG	ELAERSGHIG	LGKIQSHQLG	NLMIQQAAIK
451	GNIGYIVRFS	DHGHEVHSPF	DNHASHSDSD	EAGSPVDGFS	LYRIHWDGYE
501	HHPADGYDGP	QGGGYPAPKG	ARDIYSYDIK	GVAQNIRLNL	TDNRSTGQRL
551	ADRFHNAGSM	LTQGVGDGFK	RATRYSPELD	RSGNAAEAFN	GTADIVKNII
601	GAAGEIVGAG	DAVQGISEGS	NIAVMHGLGL	LSTENKMARI	NDLADMAQLK
651	DYAAAAIRDW	AVQNPNAAQG	IEAVSNIFMA	AIPIKGIGAV	RGKYGLGGIT
701	AHPIKRSQMG	AIALPKGKSA	VSDNFADAAY	AKYPSPYHSR	NIRSNLEQRY
751	GKENITSSTV	PPSNGKNVKL	ADQRHPKTGV	PFDGKGFPNF	EKHVKYDTLE
801	HHHHHH*				

961-741

ATGGCCACAAACGACGACGATGTTAAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATC
AACGGTTTCAAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCC
GATGTTGAAGCCGACGACTTTAAAGGTCTGGGTCTGAAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAAC
AAACAAAACGTCGATGCCAAAGTAAAAGCTGCAGAATCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGAT
GCCGCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAATTGGGAGAAAATATAACGACA
TTTGCTGAAGAGACTAAGACAAATATCGTAAAAATTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCAT
GCCGAAGCATTCAACGATATCGCCGATTCATTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAAT
GAAGCCAAACAGACGGCCGAAGAAACCAAACAAAACGTCGATGCCAAAGTAAAAGCTGCAGAAACTGCAGCAGGCAAA
GCCGAAGCTGCCGCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTCGCTGCAAAAGTTACCGACATCAAA
GCTGATATCGCTACGAACAAAGATAATATTGCTAAAAAAGCAAACAGTGCCGACGTGTACACCAGAGAAGAGTCTGAC
AGCAAATTTGTCAGAATTGATGGTCTGAACGCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCC
ATTGCCGATCACGATACTCGCCTGAACGGTTTGGATAAAACAGTGTCAGACCTGCGCAAAGAAACCCGCCAAGGCCTT

GCAGAACAAGCCGCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTCGGTTCAATGTAACGGCTGCAGTCGGCGGC
TACAAATCCGAATCGGCAGTCGCCATCGGTACCGGCTTCCGCTTTACCGAAAACTTTGCCGCCAAAGCAGGCGTGGCA
GTCGGCACTTCGTCCGGTTCTTCCGCAGCCTACCATGTCGGCGTCAATTACGAGTGGGGATCCGGAGGGGGTGGTGTC
GCCGCCGACATCGGTGCGGGGCTTGCCGATGCACTAACCGCACCGCTCGACCATAAAGACAAAGGTTTGCAGTCTTTG
ACGCTGGATCAGTCCGTCAGGAAAAACGAGAAACTGAAGCTGGCGGCACAAGGTGCGGAAAAAACTTATGGAAACGGT
GACAGCCTCAATACGGGCAAATTGAAGAACGACAAGGTCAGCCGTTTCGACTTTATCCGCCAAATCGAAGTGGACGGG
CAGCTCATTACCTTGGAGAGTGGAGAGTTCCAAGTATACAAACAAAGCCATTCCGCCTTAACCGCCTTTCAGACCGAG
CAAATACAAGATTCGGAGCATTCCGGGAAGATGGTTGCGAAACGCCAGTTCAGAATCGGCGACATAGCGGGCGAACAT
ACATCTTTTGACAAGCTTCCCGAAGGCGGCAGGGCGACATATCGCGGGACGGCGTTCGGTTCAGACGATGCCGGCGGA
AAACTGACCTACACCATAGATTTCGCCGCCAAGCAGGGAAACGGCAAAATCGAACATTTGAAATCGCCAGAACTCAAT
GTCGACCTGGCCGCCGCCGATATCAAGCCGGATGGAAAACGCCATGCCGTCATCAGCGGTTCCGTCCTTTACAACCAA
GCCGAGAAAGGCAGTTACTCCCTCGGTATCTTTGGCGGAAAAGCCCAGGAAGTTGCCGGCAGCGCGGAAGTGAAAACC
GTAAACGGCATACGCCATATCGGCCTTGCCGCCAAGCAACTCGAGCACCACCACCACCACCACTGA

```
  1  MATNDDDVKK AATVAIAAAY NNGQEINGFK AGETIYDIDE DGTITKKDAT
 51  AADVEADDFK GLGLKKVVTN LTKTVNENKQ NVDAKVKAAE SEIEKLTTKL
101  ADTDAALADT DAALDATTNA LNKLGENITT FAEETKTNIV KIDEKLEAVA
151  DTVDKHAEAF NDIADSLDET NTKADEAVKT ANEAKQTAEE TKQNVDAKVK
201  AAETAAGKAE AAAGTANTAA DKAEAVAAKV TDIKADIATN KDNIAKKANS
251  ADVYTREESD SKFVRIDGLN ATTEKLDTRL ASAEKSIADH DTRLNGLDKT
301  VSDLRKETRQ GLAEQAALSG LFQPYNVGRF NVTAAVGGYK SESAVAIGTG
351  FRFTENFAAK AGVAVGTSSG SSAAYHVGVN YEWGSGGGGV AADIGAGLAD
401  ALTAPLDHKD KGLQSLTLDQ SVRKNEKLKL AAQGAEKTYG NGDSLNTGKL
451  KNDKVSRFDF IRQIEVDGQL ITLESGEFQV YKQSHSALTA FQTEQIQDSE
501  HSGKMVAKRQ FRIGDIAGEH TSFDKLPEGG RATYRGTAFG SDDAGGKLTY
551  TIDFAAKQGN GKIEHLKSPE LNVDLAAADI KPDGKRHAVI SGSVLYNQAE
601  KGSYSLGIFG GKAQEVAGSA EVKTVNGIRH IGLAAKQLEH HHHHH*
```

961-983

ATGGCCACAAACGACGACGATGTTAAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATC
AACGGTTTCAAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCC
GATGTTGAAGCCGACGACTTTAAAGGTCTGGGTCTGAAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAAC
AAACAAAACGTCGATGCCAAAGTAAAAGCTGCAGAATCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGAT
GCCGCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAATTGGGAGAAAATATAACGACA
TTTGCTGAAGAGACTAAGACAAATATCGTAAAAATTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCAT
GCCGAAGCATTCAACGATATCGCCGATTCATTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAAT
GAAGCCAAACAGACGGCCGAAGAAACCAAACAAAACGTCGATGCCAAAGTAAAAGCTGCAGAAACTGCAGCAGGCAAA
GCCGAAGCTGCCGCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTCGCTGCAAAAGTTACCGACATCAAA
GCTGATATCGCTACGAACAAAGATAATATTGCTAAAAAAGCAAACAGTGCCGACGTGTACACCAGAGAAGAGTCTGAC
AGCAAATTTGTCAGAATTGATGGTCTGAACGCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCC
ATTGCCGATCACGATACTCGCCTGAACGGTTTGGATAAAACAGTGTCAGACCTGCGCAAAGAAACCCGCCAAGGCCTT
GCAGAACAAGCCGCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTCGGTTCAATGTAACGGCTGCAGTCGGCGGC
TACAAATCCGAATCGGCAGTCGCCATCGGTACCGGCTTCCGCTTTACCGAAAACTTTGCCGCCAAAGCAGGCGTGGCA
GTCGGCACTTCGTCCGGTTCTTCCGCAGCCTACCATGTCGGCGTCAATTACGAGTGGGGATCCGGCGGAGGCGGCACT
TCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCAGTA
TCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACA
GACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAG
AATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGACACAGGC
GAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTAT
ACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTT
ATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCCCATATTATT
GGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACGAATGAT
GAAACCAAGAACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTC
AATAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTACCGC
CAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAAC
CTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACA
TATGCCCTATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAA
AAGTTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCATTGCGGAATTACTGCC
ATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCC
TTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGT
ACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCG
GGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCC
TACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAAC
AACACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGCGTC
GAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGCGACGGCATTGTCTATCTGGCA
GATACCGACCAATCCGGCGCAAACGAAACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTAC

ACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAG
GGGGCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGGCAGGATTATTCTTTC
TTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGAC
ACGCTGTCCTATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTG
AAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACA
CCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCA
GCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGAC
AGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCACGGGT
CTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGT
ACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACA
TGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATC
GGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAACATGCG
GAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGAT
TTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGG
AGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGCGATAAAGCC
GTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCG
ACTGCAGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAA
TTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGAGTC
GGCGTAGGCTACCGGTTCCTCGAGCACCACCACCACCACCACTGA

1	MATNDDDVKK	AATVAIAAAY	NNGQEINGFK	AGETIYDIDE	DGTITKKDAT
51	AADVEADDFK	GLGLKKVVTN	LTKTVNENKQ	NVDAKVKAAE	SEIEKLTTKL
101	ADTDAALADT	DAALDATTNA	LNKLGENITT	FAEETKTNIV	KIDEKLEAVA
151	DTVDKHAEAF	NDIADSLDET	NTKADEAVKT	ANEAKQTAEE	TKQNVDAKVK
201	AAETAAGKAE	AAAGTANTAA	DKAEAVAAKV	TDIKADIATN	KDNIAKKANS
251	ADVYTREESD	SKFVRIDGLN	ATTEKLDTRL	ASAEKSIADH	DTRLNGLDKT
301	VSDLRKETRQ	GLAEQAALSG	LFQPYNVGRF	NVTAAVGGYK	SESAVAIGTG
351	FRFTENFAAK	AGVAVGTSSG	SSAAYHVGVN	YEWGSGGGGT	SAPDFNAGGT
401	GIGSNSRATT	AKSAAVSYAG	IKNEMCKDRS	MLCAGRDDVA	VTDRDAKINA
451	PPPNLHTGDF	PNPNDAYKNL	INLKPAIEAG	YTGRGVEVGI	VDTGESVGSI
501	SFPELYGRKE	HGYNENYKNY	TAYMRKEAPE	DGGGKDIEAS	FDDEAVIETE
551	AKPTDIRHVK	EIGHIDLVSH	IIGGRSVDGR	PAGGIAPDAT	LHIMNTNDET
601	KNEMMVAAIR	NAWVKLGERG	VRIVNNSFGT	TSRAGTADLF	QIANSEEQYR
651	QALLDYSGGD	KTDEGIRLMQ	QSDYGNLSYH	IRNKNMLFIF	STGNDAQAQP
701	NTYALLPFYE	KDAQKGIITV	AGVDRSGEKF	KREMYGEPGT	EPLEYGSNHC
751	GITAMWCLSA	PYEASVRFTR	TNPIQIAGTS	FSAPIVTGTA	ALLLQKYPWM
801	SNDNLRTTLL	TTAQDIGAVG	VDSKFGWGLL	DAGKAMNGPA	SFPFGDFTAD
851	TKGTSDIAYS	FRNDISGTGG	LIKKGGSQLQ	LHGNNTYTGK	TIIEGGSLVL
901	YGNNKSDMRV	ETKGALIYNG	AASGGSLNSD	GIVYLADTDQ	SGANETVHIK
951	GSLQLDGKGT	LYTRLGKLLK	VDGTAIIGGK	LYMSARGKGA	GYLNSTGRRV
1001	PFLSAAKIGQ	DYSFFTNIET	DGGLLASLDS	VEKTAGSEGD	TLSYYVRRGN
1051	AARTASAAAH	SAPAGLKHAV	EQGGSNLENL	MVELDASESS	ATPETVETAA
1101	ADRTDMPGIR	PYGATFRAAA	AVQHANAADG	VRIFNSLAAT	VYADSTAAHA
1151	DMQGRRLKAV	SDGLDHNGTG	LRVIAQTQQD	GGTWEQGGVE	GKMRGSTQTV
1201	GIAAKTGENT	TAAATLGMGR	STWSENSANA	KTDSISLFAG	IRHDAGDIGY
1251	LKGLFSYGRY	KNSISRSTGA	DEHAEGSVNG	TLMQLGALGG	VNVPFAATGD
1301	LTVEGGLRYD	LLKQDAFAEK	GSALGWSGNS	LTEGTLVGLA	GLKLSQPLSD
1351	KAVLFATAGV	ERDLNGRDYT	VTGGFTGATA	ATGKTGARNM	PHTRLVAGLG
1401	ADVEFGNGWN	GLARYSYAGS	KQYGNHSGRV	GVGYRFLEHH	HHHH*

961c-ORF46.1

ATGGCCACAAACGACGACGATGTTAAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATC
AACGGTTTCAAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCC
GATGTTGAAGCCGACGACTTTAAAGGTCTGGGTCTGAAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAAC
AAACAAAACGTCGATGCCAAAGTAAAAGCTGCAGAATCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGAT
GCCGCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAATTGGGAGAAAATATAACGACA
TTTGCTGAAGAGACTAAGACAAATATCGTAAAAATTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCAT
GCCGAAGCATTCAACGATATCGCCGATTCATTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAAT
GAAGCCAAACAGACGGCCGAAGAAACCAAACAAAACGTCGATGCCAAAGTAAAAGCTGCAGAAACTGCAGCAGGCAAA
GCCGAAGCTGCCGCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTCGCTGCAAAAGTTACCGACATCAAA
GCTGATATCGCTACGAACAAAGATAATATTGCTAAAAAAGCAAACAGTGCCGACGTGTACACCAGAGAAGAGTCTGAC
AGCAAATTTGTCAGAATTGATGGTCTGAACGCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCC
ATTGCCGATCACGATACTCGCCTGAACGGTTTGGATAAAACAGTGTCAGACCTGCGCAAAGAAACCCGCCAAGGCCTT
GCAGAACAAGCCGCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTGGATCCGGAGGAGGAGGATCAGATTTGGCA
AACGATTCTTTTATCCGGCAGGTTCTCGACCGTCAGCATTTCGAACCCGACGGGAAATACCACCTATTCGGCAGCAGG
GGGGAACTTGCCGAGCGCAGCGGCCATATCGGATTGGGAAAAATACAAAGCCATCAGTTGGGCAACCTGATGATTCAA

CAGGCGGCCATTAAAGGAAATATCGGCTACATTGTCCGCTTTTCCGATCACGGGCACGAAGTCCATTCCCCCTTCGAC
AACCATGCCTCACATTCCGATTCTGATGAAGCCGGTAGTCCCGTTGACGGATTTAGCCTTTACCGCATCCATTGGGAC
GGATACGAACACCATCCCGCCGACGGCTATGACGGCCACAGGGCGGCGGCTATCCCGCTCCCAAAGGCGCGAGGGAT
ATATACAGCTACGACATAAAAGGCGTTGCCCAAAATATCCGCCTCAACCTGACCGACAACCGCAGCACCGGACAACGG
CTTGCCGACCGTTTCCACAATGCCGGTAGTATGCTGACGCAAGGAGTAGGCGACGGATTCAAACGCGCCACCCGATAC
AGCCCCGAGCTGGACAGATCGGGCAATGCCGCCGAAGCCTTCAACGGCACTGCAGATATCGTTAAAAACATCATCGGC
GCGGCAGGAGAAATTGTCGGCGCAGGCGATGCCGTGCAGGGCATAAGCGAAGGCTCAAACATTGCTGTCATGCACGGC
TTGGGTCTGCTTTCCACCGAAAACAAGATGGCGCGCATCAACGATTTGGCAGATATGGCGCAACTCAAAGACTATGCC
GCAGCAGCCATCCGCGATTGGGCAGTCCAAAACCCCAATGCCGCACAAGGCATAGAAGCCGTCAGCAATATCTTTATG
GCAGCCATCCCCATCAAAGGGATTGGAGCTGTTCGGGGAAAATACGGCTTGGGCGGCATCACGGCACATCCTATCAAG
CGGTCGCAGATGGGCGCGATCGCATTGCCGAAAGGGAAATCCGCCGTCAGCGACAATTTTGCCGATGCGGCATACGCC
AAATACCCGTCCCCTTACCATTCCCGAAATATCCGTTCAAACTTGGAGCAGCGTTACGGCAAAGAAAACATCACCTCC
TCAACCGTGCCGCCGTCAAACGGCAAAAATGTCAAACTGGCAGACCAACGCCACCCGAAGACAGGCGTACCGTTTGAC
GGTAAAGGGTTTCCGAATTTTGAGAAGCACGTGAAATATGATACGCTCGAGCACCACCACCACCACCACTGA

1 MATNDDDVKK AATVAIAAAY NNGQEINGFK AGETIYDIDE DGTITKKDAT
51 AADVEADDFK GLGLKKVVTN LTKTVNENKQ NVDAKVKAAE SEIEKLTTKL
101 ADTDAALADT DAALDATTNA LNKLGENITT FAEETKTNIV KIDEKLEAVA
151 DTVDKHAEAF NDIADSLDET NTKADEAVKT ANEAKQTAEE TKQNVDAKVK
201 AAETAAGKAE AAAGTANTAA DKAEAVAAKV TDIKADIATN KDNIAKKANS
251 ADVYTREESD SKFVRIDGLN ATTEKLDTRL ASAEKSIADH DTRLNGLDKT
301 VSDLRKETRQ GLAEQAALSG LFQPYNVGGS GGGGSDLAND SFIRQVLDRQ
351 HFEPDGKYHL FGSRGELAER SGHIGLGKIQ SHQLGNLMIQ QAAIKGNIGY
401 IVRFSDHGHE VHSPFDNHAS HSDSDEAGSP VDGFSLYRIH WDGYEHHPAD
451 GYDGPQGGGY PAPKGARDIY SYDIKGVAQN IRLNLTDNRS TGQRLADRFH
501 NAGSMLTQGV GDGFKRATRY SPELDRSGNA AEAFNGTADI VKNIIGAAGE
551 IVGAGDAVQG ISEGSNIAVM HGLGLLSTEN KMARINDLAD MAQLKDYAAA
601 AIRDWAVQNP NAAQGIEAVS NIFMAAIPIK GIGAVRGKYG LGGITAHPIK
651 RSQMGAIALP KGKSAVSDNF ADAAYAKYPS PYHSRNIRSN LEQRYGKENI
701 TSSTVPPSNG KNVKLADQRH PKTGVPFDGK GFPNFEKHVK YDTLEHHHHH
751 H*

961c-741

ATGGCCACAAACGACGACGATGTTAAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATC
AACGGTTTCAAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCC
GATGTTGAAGCCGACGACTTTAAAGGTCTGGGTCTGAAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAAC
AAACAAAACGTCGATGCCAAAGTAAAAGCTGCAGAATCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGAT
GCCGCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAATTGGGAGAAAATATAACGACA
TTTGCTGAAGAGACTAAGACAAATATCGTAAAAATTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCAT
GCCGAAGCATTCAACGATATCGCCGATTCATTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAAT
GAAGCCAAACAGACGGCCGAAGAAACCAAACAAAACGTCGATGCCAAAGTAAAAGCTGCAGAAACTGCAGCAGGCAAA
GCCGAAGCTGCCGCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTCGCTGCAAAAGTTACCGACATCAAA
GCTGATATCGCTACGAACAAAGATAATATTGCTAAAAAAGCAAACAGTGCCGACGTGTACACCAGAGAAGAGTCTGAC
AGCAAATTTGTCAGAATTGATGGTCTGAACGCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCC
ATTGCCGATCACGATACTCGCCTGAACGGTTTGGATAAAACAGTGTCAGACCTGCGCAAAGAAACCCGCCAAGGCCTT
GCAGAACAAGCCGCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTGGATCCGGAGGGGGTGGTGTCGCCGCCGAC
ATCGGTGCGGGGCTTGCCGATGCACTAACCGCACCGCTCGACCATAAAGACAAAGGTTTGCAGTCTTTGACGCTGGAT
CAGTCCGTCAGGAAAAACGAGAAACTGAAGCTGGCGGCACAAGGTGCGGAAAAAACTTATGGAAACGGTGACAGCCTC
AATACGGGCAAATTGAAGAACGACAAGGTCAGCCGTTTCGACTTTATCCGCCAAATCGAAGTGGACGGGCAGCTCATT
ACCTTGGAGAGTGGAGAGTTCCAAGTATACAAACAAAGCCATTCCGCCTTAACCGCCTTTCAGACCGAGCAAATACAA
GATTCGGAGCATTCCGGGAAGATGGTTGCGAAACGCCAGTTCAGAATCGGCGACATAGCGGGCGAACATACATCTTTT
GACAAGCTTCCCGAAGGCGGCAGGGCGACATATCGCGGGACGGCGTTCGGTTCAGACGATGCCGGCGGAAAACTGACC
TACACCATAGATTTCGCCGCCAAGCAGGGAAACGGCAAAATCGAACATTTGAAATCGCCAGAACTCAATGTCGACCTG
GCCGCCGCCGATATCAAGCCGGATGGAAAACGCCATGCCGTCATCAGCGGTTCCGTCCTTTACAACCAAGCCGAGAAA
GGCAGTTACTCCCTCGGTATCTTTGGCGGAAAAGCCCAGGAAGTTGCCGGCAGCGCGGAAGTGAAAACCGTAAACGGC
ATACGCCATATCGGCCTTGCCGCCAAGCAACTCGAGCACCACCACCACCACCACTGA

1 MATNDDDVKK AATVAIAAAY NNGQEINGFK AGETIYDIDE DGTITKKDAT
51 AADVEADDFK GLGLKKVVTN LTKTVNENKQ NVDAKVKAAE SEIEKLTTKL
101 ADTDAALADT DAALDATTNA LNKLGENITT FAEETKTNIV KIDEKLEAVA
151 DTVDKHAEAF NDIADSLDET NTKADEAVKT ANEAKQTAEE TKQNVDAKVK
201 AAETAAGKAE AAAGTANTAA DKAEAVAAKV TDIKADIATN KDNIAKKANS
251 ADVYTREESD SKFVRIDGLN ATTEKLDTRL ASAEKSIADH DTRLNGLDKT
301 VSDLRKETRQ GLAEQAALSG LFQPYNVGGS GGGGVAADIG AGLADALTAP
351 LDHKDKGLQS LTLDQSVRKN EKLKLAAQGA EKTYGNGDSL NTGKLKNDKV
401 SRFDFIRQIE VDGQLITLES GEFQVYKQSH SALTAFQTEQ IQDSEHSGKM

```
451     VAKRQFRIGD IAGEHTSFDK LPEGGRATYR GTAFGSDDAG GKLTYTIDFA
501     AKQGNGKIEH LKSPELNVDL AAADIKPDGK RHAVISGSVL YNQAEKGSYS
551     LGIFGGKAQE VAGSAEVKTV NGIRHIGLAA KQLEHHHHHH *
```

961c-983

```
ATGGCCACAAACGACGACGATGTTAAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATC
AACGGTTTCAAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCC
GATGTTGAAGCCGACGACTTTAAAGGTCTGGGTCTGAAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAAC
AAACAAAACGTCGATGCCAAAGTAAAAGCTGCAGAATCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGAT
GCCGCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAATTGGGAGAAAATATAACGACA
TTTGCTGAAGAGACTAAGACAAATATCGTAAAAATTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCAT
GCCGAAGCATTCAACGATATCGCCGATTCATTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAAT
GAAGCCAAACAGACGGCCGAAGAAACCAAACAAAACGTCGATGCCAAAGTAAAAGCTGCAGAAACTGCAGCAGGCAAA
GCCGAAGCTGCCGCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTCGCTGCAAAAGTTACCGACATCAAA
GCTGATATCGCTACGAACAAAGATAATATTGCTAAAAAAGCAAACAGTGCCGACGTGTACACCAGAGAAGAGTCTGAC
AGCAAATTTGTCAGAATTGATGGTCTGAACGCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCC
ATTGCCGATCACGATACTCGCCTGAACGGTTTGGATAAAACAGTGTCAGACCTGCGCAAAGAAACCCGCCAAGGCCTT
GCAGAACAAGCCGCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTGGATCCGGCGGAGGCGGCACTTCTGCGCCC
GACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCC
GGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGAT
GCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAATTTGATC
AACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGACACAGGCGAATCCGTC
GGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTAT
ATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT
GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCCCATATTATTGGCGGGCGT
TCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACGAATGATGAAACCAAG
AACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGT
TTTGGAACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTG
CTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACCTGTCCTAC
CACATCCGTAATAAAAACATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCTA
TTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAAA
CGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGC
CTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCA
CCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTG
CTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCC
ATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTC
CGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTAT
ACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAAA
GGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGAC
CAATCCGGCGCAAACGAAACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTTG
GGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAGGGGGCAGGC
TATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGGCAGGATTATTCTTTCTTCACAAAC
ATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC
TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC
GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACG
GTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCC
GTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCC
GCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCACGGGTCTGCGCGTC
ATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACC
GTCGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGACGCAGCACATGGAGCGAA
AACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCTATCTC
AAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGC
GTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC
GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC
AGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTT
GCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCA
ACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATTCGGCAAC
GGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGAGTCGGCGTAGGC
TACCGGTTCCTCGAGCACCACCACCACCACCACTGA
```

```
  1     MATNDDDVKK AATVAIAAAY NNGQEINGFK AGETIYDIDE DGTITKKDAT
 51     AADVEADDFK GLGLKKVVTN LTKTVNENKQ NVDAKVKAAE SEIEKLTTKL
101     ADTDAALADT DAALDATTNA LNKLGENITT FAEETKTNIV KIDEKLEAVA
151     DTVDKHAEAF NDIADSLDET NTKADEAVKT ANEAKQTAEE TKQNVDAKVK
201     AAETAAGKAE AAAGTANTAA DKAEAVAAKV TDIKADIATN KDNIAKKANS
251     ADVYTREESD SKFVRIDGLN ATTEKLDTRL ASAEKSIADH DTRLNGLDKT
```

```
 301  VSDLRKETRQ GLAEQAALSG LFQPYNVGGS GGGGTSAPDF NAGGTGIGSN
 351  SRATTAKSAA VSYAGIKNEM CKDRSMLCAG RDDVAVTDRD AKINAPPPNL
 401  HTGDFPNPND AYKNLINLKP AIEAGYTGRG VEVGIVDTGE SVGSISFPEL
 451  YGRKEHGYNE NYKNYTAYMR KEAPEDGGGK DIEASFDDEA VIETEAKPTD
 501  IRHVKEIGHI DLVSHIIGGR SVDGRPAGGI APDATLHIMN TNDETKNEMM
 551  VAAIRNAWVK LGERGVRIVN NSFGTTSRAG TADLFQIANS EEQYRQALLD
 601  YSGGDKTDEG IRLMQQSDYG NLSYHIRNKN MLFIFSTGND AQAQPNTYAL
 651  LPFYEKDAQK GIITVAGVDR SGEKFKREMY GEPGTEPLEY GSNHCGITAM
 701  WCLSAPYEAS VRFTRTNPIQ IAGTSFSAPI VTGTAALLLQ KYPWMSNDNL
 751  RTTLLTTAQD IGAVGVDSKF GWGLLDAGKA MNGPASFPFG DFTADTKGTS
 801  DIAYSFRNDI SGTGGLIKKG GSQLQLHGNN TYTGKTIIEG GSLVLYGNNK
 851  SDMRVETKGA LIYNGAASGG SLNSDGIVYL ADTDQSGANE TVHIKGSLQL
 901  DGKGTLYTRL GKLLKVDGTA IIGGKLYMSA RGKGAGYLNS TGRRVPFLSA
 951  AKIGQDYSFF TNIETDGGLL ASLDSVEKTA GSEGDTLSYY VRRGNAARTA
1001  SAAAHSAPAG LKHAVEQGGS NLENLMVELD ASESSATPET VETAAADRTD
1051  MPGIRPYGAT FRAAAAVQHA NAADGVRIFN SLAATVYADS TAAHADMQGR
1101  RLKAVSDGLD HNGTGLRVIA QTQQDGGTWE QGGVEGKMRG STQTVGIAAK
1151  TGENTTAAAT LGMGRSTWSE NSANAKTDSI SLFAGIRHDA GDIGYLKGLF
1201  SYGRYKNSIS RSTGADEHAE GSVNGTLMQL GALGGVNVPF AATGDLTVEG
1251  GLRYDLLKQD AFAEKGSALG WSGNSLTEGT LVGLAGLKLS QPLSDKAVLF
1301  ATAGVERDLN GRDYTVTGGF TGATAATGKT GARNMPHTRL VAGLGADVEF
1351  GNGWNGLARY SYAGSKQYGN HSGRVGVGYR FLEHHHHHH*
```

961cL-ORF46.1

```
ATGAAACACTTTCCATCCAAAGTACTGACCACAGCCATCCTTGCCACTTTCTGTAGCGGCGCACTGGCAGCCACAAAC
GACGACGATGTTAAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATCAACGGTTTCAAA
GCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCCGATGTTGAAGCC
GACGACTTTAAAGGTCTGGGTCTGAAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAACAAACAAAACGTC
GATGCCAAAGTAAAAGCTGCAGAATCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGATGCCGCTTTAGCA
GATACTGATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAATTGGGAGAAAATATAACGACATTTGCTGAAGAG
ACTAAGACAAATATCGTAAAAATTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCATGCCGAAGCATTC
AACGATATCGCCGATTCATTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAATGAAGCCAAACAG
ACGGCCGAAGAAACCAAACAAAACGTCGATGCCAAAGTAAAAGCTGCAGAAACTGCAGCAGGCAAAGCCGAAGCTGCC
GCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTCGCTGCAAAAGTTACCGACATCAAAGCTGATATCGCT
ACGAACAAAGATAATATTGCTAAAAAAGCAAACAGTGCCGACGTGTACACCAGAGAAGAGTCTGACAGCAAATTTGTC
AGAATTGATGGTCTGAACGCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCCATTGCCGATCAC
GATACTCGCCTGAACGGTTTGGATAAAACAGTGTCAGACCTGCGCAAAGAAACCCGCCAAGGCCTTGCAGAACAAGCC
GCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTGGATCCGGAGGAGGAGGATCAGATTTGGCAAACGATTCTTTT
ATCCGGCAGGTTCTCGACCGTCAGCATTTCGAACCCGACGGGAAATACCACCTATTCGGCAGCAGGGGGGAACTTGCC
GAGCGCAGCGGCCATATCGGATTGGGAAAAATACAAAGCCATCAGTTGGGCAACCTGATGATTCAACAGGCGGCCATT
AAAGGAAATATCGGCTACATTGTCCGCTTTTCCGATCACGGGCACGAAGTCCATTCCCCCTTCGACAACCATGCCTCA
CATTCCGATTCTGATGAAGCCGGTAGTCCCGTTGACGGATTTAGCCTTTACCGCATCCATTGGGACGGATACGAACAC
CATCCCGCCGACGGCTATGACGGGCCACAGGGCGGCGGCTATCCCGCTCCCAAAGGCGCGAGGGATATATACAGCTAC
GACATAAAAGGCGTTGCCCAAAATATCCGCCTCAACCTGACCGACAACCGCAGCACCGGACAACGGCTTGCCGACCGT
TTCCACAATGCCGGTAGTATGCTGACGCAAGGAGTAGGCGACGGATTCAAACGCGCCACCCGATACAGCCCCGAGCTG
GACAGATCGGGCAATGCCGCCGAAGCCTTCAACGGCACTGCAGATATCGTTAAAAACATCATCGGCGCGGCAGGAGAA
ATTGTCGGCGCAGGCGATGCCGTGCAGGGCATAAGCGAAGGCTCAAACATTGCTGTCATGCACGGCTTGGGTCTGCTT
TCCACCGAAAACAAGATGGCGCGCATCAACGATTTGGCAGATATGGCGCAACTCAAAGACTATGCCGCAGCAGCCATC
CGCGATTGGGCAGTCCAAAACCCCAATGCCGCACAAGGCATAGAAGCCGTCAGCAATATCTTTATGGCAGCCATCCCC
ATCAAAGGGATTGGAGCTGTTCGGGGAAAATACGGCTTGGGCGGCATCACGGCACATCCTATCAAGCGGTCGCAGATG
GGCGCGATCGCATTGCCGAAAGGGAAATCCGCCGTCAGCGACAATTTTGCCGATGCGGCATACGCCAAATACCCGTCC
CCTTACCATTCCCGAAATATCCGTTCAAACTTGGAGCAGCGTTACGGCAAAGAAAACATCACCTCCTCAACCGTGCCG
CCGTCAAACGGCAAAAATGTCAAACTGGCAGACCAACGCCACCCGAAGACAGGCGTACCGTTTGACGGTAAAGGGTTT
CCGAATTTTGAGAAGCACGTGAAATATGATACGTAACTCGAG
```

```
  1  MKHFPSKVLT TAILATFCSG ALAATNDDDV KKAATVAIAA AYNNGQEING
 51  FKAGETIYDI DEDGTITKKD ATAADVEADD FKGLGLKKVV TNLTKTVNEN
101  KQNVDAKVKA AESEIEKLTT KLADTDAALA DTDAALDATT NALNKLGENI
151  TTFAEETKTN IVKIDEKLEA VADTVDKHAE AFNDIADSLD ETNTKADEAV
201  KTANEAKQTA EETKQNVDAK VKAAETAAGK AEAAAGTANT AADKAEAVAA
251  KVTDIKADIA TNKDNIAKKA NSADVYTREE SDSKFVRIDG LNATTEKLDT
301  RLASAEKSIA DHDTRLNGLD KTVSDLRKET RQGLAEQAAL SGLFQPYNVG
351  GSGGGGSDLA NDSFIRQVLD RQHFEPDGKY HLFGSRGELA ERSGHIGLGK
401  IQSHQLGNLM IQQAIKGNI GYIVRFSDHG HEVHSPFDNH ASHSDSDEAG
451  SPVDGFSLYR IHWDGYEHHP ADGYDGPQGG GYPAPKGARD IYSYDIKGVA
501  QNIRLNLTDN RSTGQRLADR FHNAGSMLTQ GVGDGFKRAT RYSPELDRSG
```

551 NAAEAFNGTA DIVKNIIGAA GEIVGAGDAV QGISEGSNIA VMHGLGLLST
601 ENKMARINDL ADMAQLKDYA AAAIRDWAVQ NPNAAQGIEA VSNIFMAAIP
651 IKGIGAVRGK YGLGGITAHP IKRSQMGAIA LPKGKSAVSD NFADAAYAKY
701 PSPYHSRNIR SNLEQRYGKE NITSSTVPPS NGKNVKLADQ RHPKTGVPFD
751 GKGFPNFEKH VKYDT*

961cL-741

ATGAAACACTTTCCATCCAAAGTACTGACCACAGCCATCCTTGCCACTTTCTGTAGCGGCGCACTGGCAGCCACAAAC
GACGACGATGTTAAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATCAACGGTTTCAAA
GCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCCGATGTTGAAGCC
GACGACTTTAAAGGTCTGGGTCTGAAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAACAAACAAAACGTC
GATGCCAAAGTAAAAGCTGCAGAATCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGATGCCGCTTTAGCA
GATACTGATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAATTGGGAGAAAATATAACGACATTTGCTGAAGAG
ACTAAGACAAATATCGTAAAAATTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCATGCCGAAGCATTC
AACGATATCGCCGATTCATTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAATGAAGCCAAACAG
ACGGCCGAAGAAACCAAACAAAACGTCGATGCCAAAGTAAAAGCTGCAGAAACTGCAGCAGGCAAAGCCGAAGCTGCC
GCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTCGCTGCAAAAGTTACCGACATCAAAGCTGATATCGCT
ACGAACAAAGATAATATTGCTAAAAAAGCAAACAGTGCCGACGTGTACACCAGAGAAGAGTCTGACAGCAAATTTGTC
AGAATTGATGGTCTGAACGCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCCATTGCCGATCAC
GATACTCGCCTGAACGGTTTGGATAAAACAGTGTCAGACCTGCGCAAAGAAACCCGCCAAGGCCTTGCAGAACAAGCC
GCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTGGATCCGGAGGGGGTGGTGTCGCCGCCGACATCGGTGCGGGG
CTTGCCGATGCACTAACCGCACCGCTCGACCATAAAGACAAAGGTTTGCAGTCTTTGACGCTGGATCAGTCCGTCAGG
AAAAACGAGAAACTGAAGCTGGCGGCACAAGGTGCGGAAAAAACTTATGGAAACGGTGACAGCCTCAATACGGGCAAA
TTGAAGAACGACAAGGTCAGCCGTTTCGACTTTATCCGCCAAATCGAAGTGGACGGGCAGCTCATTACCTTGGAGAGT
GGAGAGTTCCAAGTATACAAACAAAGCCATTCCGCCTTAACCGCCTTTCAGACCGAGCAAATACAAGATTCGGAGCAT
TCCGGGAAGATGGTTGCGAAACGCCAGTTCAGAATCGGCGACATAGCGGGCGAACATACATCTTTTGACAAGCTTCCC
GAAGGCGGCAGGGCGACATATCGCGGGACGGCGTTCGGTTCAGACGATGCCGGCGGAAAACTGACCTACACCATAGAT
TTCGCCGCCAAGCAGGGAAACGGCAAAATCGAACATTTGAAATCGCCAGAACTCAATGTCGACCTGGCCGCCGCCGAT
ATCAAGCCGGATGGAAAACGCCATGCCGTCATCAGCGGTTCCGTCCTTTACAACCAAGCCGAGAAAGGCAGTTACTCC
CTCGGTATCTTTGGCGGAAAAGCCCAGGAAGTTGCCGGCAGCGCGGAAGTGAAAACCGTAAACGGCATACGCCATATC
GGCCTTGCCGCCAAGCAACTCGAGCACCACCACCACCACCACTGA

1 MKHFPSKVLT TAILATFCSG ALAATNDDDV KKAATVAIAA AYNNGQEING
51 FKAGETIYDI DEDGTITKKD ATAADVEADD FKGLGLKKVV TNLTKTVNEN
101 KQNVDAKVKA AESEIEKLTT KLADTDAALA DTDAALDATT NALNKLGENI
151 TTFAEETKTN IVKIDEKLEA VADTVDKHAE AFNDIADSLD ETNTKADEAV
201 KTANEAKQTA EETKQNVDAK VKAAETAAGK AEAAAGTANT AADKAEAVAA
251 KVTDIKADIA TNKDNIAKKA NSADVYTREE SDSKFVRIDG LNATTEKLDT
301 RLASAEKSIA DHDTRLNGLD KTVSDLRKET RQGLAEQAAL SGLFQPYNVG
351 GSGGGGVAAD IGAGLADALT APLDHKDKGL QSLTLDQSVR KNEKLKLAAQ
401 GAEKTYGNGD SLNTGKLKND KVSRFDFIRQ IEVDGQLITL ESGEFQVYKQ
451 SHSALTAFQT EQIQDSEHSG KMVAKRQFRI GDIAGEHTSF DKLPEGGRAT
501 YRGTAFGSDD AGGKLTYTID FAAKQGNGKI EHLKSPELNV DLAAADIKPD
551 GKRHAVISGS VLYNQAEKGS YSLGIFGGKA QEVAGSAEVK TVNGIRHIGL
601 AAKQLEHHHH HH*

961cL-983

ATGAAACACTTTCCATCCAAAGTACTGACCACAGCCATCCTTGCCACTTTCTGTAGCGGCGCACTGGCAGCCACAAAC
GACGACGATGTTAAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATCAACGGTTTCAAA
GCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCCGATGTTGAAGCC
GACGACTTTAAAGGTCTGGGTCTGAAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAACAAACAAAACGTC
GATGCCAAAGTAAAAGCTGCAGAATCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGATGCCGCTTTAGCA
GATACTGATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAATTGGGAGAAAATATAACGACATTTGCTGAAGAG
ACTAAGACAAATATCGTAAAAATTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCATGCCGAAGCATTC
AACGATATCGCCGATTCATTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAATGAAGCCAAACAG
ACGGCCGAAGAAACCAAACAAAACGTCGATGCCAAAGTAAAAGCTGCAGAAACTGCAGCAGGCAAAGCCGAAGCTGCC
GCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTCGCTGCAAAAGTTACCGACATCAAAGCTGATATCGCT
ACGAACAAAGATAATATTGCTAAAAAAGCAAACAGTGCCGACGTGTACACCAGAGAAGAGTCTGACAGCAAATTTGTC
AGAATTGATGGTCTGAACGCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCCATTGCCGATCAC
GATACTCGCCTGAACGGTTTGGATAAAACAGTGTCAGACCTGCGCAAAGAAACCCGCCAAGGCCTTGCAGAACAAGCC
GCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTGGATCCGGCGGAGGCGGCACTTCTGCGCCCGACTTCAATGCA
GGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAAC
GAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT
GCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCT
GCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGACACAGGCGAATCCGTCGGCAGCATATCC

```
TTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAA
GCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCG
ACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGC
AGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGATG
GTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGAACAACA
TCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGACTATTCC
GGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAAT
AAAAACATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTAT
GAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAAACGGGAAATGTAT
GGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCC
TATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACC
GGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTGCTGACGACGGCT
CAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCC
GCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATT
TCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGCAAAACC
ATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATT
TATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCA
AACGAAACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG
AAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAGGGGGCAGGCTATCTCAACAGT
ACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGGCAGGATTATTCTTTCTTCACAAACATCGAAACCGAC
GGCGGCCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGT
CGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGC
GGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCG
GCAGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCG
AATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGAT
ATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACC
CAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC
GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGAAAACAGTGCAAAT
GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTC
TCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACG
CTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTG
CGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAACAGCCTCACTGAA
GGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGC
GTGGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACG
GGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATTCGGCAACGGCTGGAACGGC
TTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA
CTCGAG
```

```
   1   MKHFPSKVLT TAILATFCSG ALAATNDDDV KKAATVAIAA AYNNGQEING
  51   FKAGETIYDI DEDGTITKKD ATAADVEADD FKGLGLKKVV TNLTKTVNEN
 101   KQNVDAKVKA AESEIEKLTT KLADTDAALA DTDAALDATT NALNKLGENI
 151   TTFAEETKTN IVKIDEKLEA VADTVDKHAE AFNDIADSLD ETNTKADEAV
 201   KTANEAKQTA EETKQNVDAK VKAAETAAGK AEAAAGTANT AADKAEAVAA
 251   KVTDIKADIA TNKDNIAKKA NSADVYTREE SDSKFVRIDG LNATTEKLDT
 301   RLASAEKSIA DHDTRLNGLD KTVSDLRKET RQGLAEQAAL SGLFQPYNVG
 351   GSGGGGTSAP DFNAGGTGIG SNSRATTAKS AAVSYAGIKN EMCKDRSMLC
 401   AGRDDVAVTD RDAKINAPPP NLHTGDFPNP NDAYKNLINL KPAIEAGYTG
 451   RGVEVGIVDT GESVGSISFP ELYGRKEHGY NENYKNYTAY MRKEAPEDGG
 501   GKDIEASFDD EAVIETEAKP TDIRHVKEIG HIDLVSHIIG GRSVDGRPAG
 551   GIAPDATLHI MNTNDETKNE MMVAAIRNAW VKLGERGVRI VNNSFGTTSR
 601   AGTADLFQIA NSEEQYRQAL LDYSGGDKTD EGIRLMQQSD YGNLSYHIRN
 651   KNMLFIFSTG NDAQAQPNTY ALLPFYEKDA QKGIITVAGV DRSGEKFKRE
 701   MYGEPGTEPL EYGSNHCGIT AMWCLSAPYE ASVRFTRTNP IQIAGTSFSA
 751   PIVTGTAALL LQKYPWMSND NLRTTLLTTA QDIGAVGVDS KFGWGLLDAG
 801   KAMNGPASFP FGDFTADTKG TSDIAYSFRN DISGTGGLIK KGGSQLQLHG
 851   NNTYTGKTII EGGSLVLYGN NKSDMRVETK GALIYNGAAS GGSLNSDGIV
 901   YLADTDQSGA NETVHIKGSL QLDGKGTLYT RLGKLLKVDG TAIIGGKLYM
 951   SARGKGAGYL NSTGRRVPFL SAAKIGQDYS FFTNIETDGG LLASLDSVEK
1001   TAGSEGDTLS YYVRRGNAAR TASAAAHSAP AGLKHAVEQG GSNLENLMVE
1051   LDASESSATP ETVETAAADR TDMPGIRPYG ATFRAAAAVQ HANAADGVRI
1101   FNSLAATVYA DSTAAHADMQ GRRLKAVSDG LDHNGTGLRV IAQTQQDGGT
1151   WEQGGVEGKM RGSTQTVGIA AKTGENTTAA ATLGMGRSTW SENSANAKTD
1201   SISLFAGIRH DAGDIGYLKG LFSYGRYKNS ISRSTGADEH AEGSVNGTLM
1251   QLGALGGVNV PFAATGDLTV EGGLRYDLLK QDAFAEKGSA LGWSGNSLTE
1301   GTLVGLAGLK LSQPLSDKAV LFATAGVERD LNGRDYTVTG GFTGATAATG
1351   KTGARNMPHT RLVAGLGADV EFGNGWNGLA RYSYAGSKQY GNHSGRVGVG
1401   YRF*
```

It will be understood that the invention has been described by way of example only and modifications may be made whilst remaining within the scope and spirit of the invention. For instance, the use of proteins from other strains is envisaged [*e.g.* see WO00/66741 for polymorphic sequences for ORF4, ORF40, ORF46, 225, 235, 287, 519, 726, 919 and 953].

EXPERIMENTAL DETAILS

Cloning strategy and oligonucleotide design

Genes coding for antigens of interest were amplified by PCR, using oligonucleotides designed on the basis of the genomic sequence of *N. meningitidis* B MC58. Genomic DNA from strain 2996 was always used as a template in PCR reactions, unless otherwise specified, and the amplified fragments were cloned in the expression vector pET21b+ (Novagen) to express the protein as C-terminal His-tagged product, or in pET-24b+(Novagen) to express the protein in 'untagged' form (*e.g.* ΔG 287K).

Where a protein was expressed without a fusion partner and with its own leader peptide (if present), amplification of the open reading frame (ATG to STOP codons) was performed.

Where a protein was expressed in 'untagged' form, the leader peptide was omitted by designing the 5'-end amplification primer downstream from the predicted leader sequence.

The melting temperature of the primers used in PCR depended on the number and type of hybridising nucleotides in the whole primer, and was determined using the formulae:

$T_{m1} = 4\,(G+C) + 2\,(A+T)$ (tail excluded)

$T_{m2} = 64.9 + 0.41\,(\%\ GC) - 600/N$ (whole primer)

The melting temperatures of the selected oligonucleotides were usually 65-70°C for the whole oligo and 50-60°C for the hybridising region alone.

Oligonucleotides were synthesised using a Perkin Elmer 394 DNA/RNA Synthesizer, eluted from the columns in 2.0ml NH_4OH, and deprotected by 5 hours incubation at 56°C. The oligos were precipitated by addition of 0.3M Na-Acetate and 2 volumes ethanol. The samples were centrifuged and the pellets resuspended in water.

		Sequences	Restriction site
fu (961)-	Fwd	CGC<u>GGATCC</u> -GGAGGGGGTGGTGTCG	BamHI

741(MC58)-His			
	Rev	CCCG<u>CTCGAG</u>-TTGCTTGGCGGCAAGGC	XhoI
fu (961)-983-His	Fwd	CGC<u>GGATCC</u> - GGCGGAGGCGGCACTT	BamHI
	Rev	CCCG<u>CTCGAG</u>-GAACCGGTAGCCTACG	XhoI
fu (961)- Orf46.1-His	Fwd	CGC<u>GGATCC</u>GGTGGTGGTGGT-TCAGATTTGGCAAACGATTC	BamHI
	Rev	CCCG<u>CTCGAG</u>-CGTATCATATTTCACGTGC	XhoI
fu (961 c-L)-741(MC58)	Fwd	CGC<u>GGATCC</u> -GGAGGGGGTGGTGTCG	BamHI
	Rev	CCCG<u>CTCGAG</u>-TTATTGCTTGGCGGCAAG	XhoI
fu (961c-L)-983	Fwd	CGC<u>GGATCC</u> - GGCGGAGGCGGCACTT	BamHI
	Rev	CCCG<u>CTCGAG</u>-TCAGAACCGGTAGCCTAC	XhoI
fu (961c-L)-Orf46.1	Fwd	CGC<u>GGATCC</u>GGTGGTGGTGGT-TCAGATTTGGCAAACGATTC	BamHI
	Rev	CCCG<u>CTCGAG</u>-TTACGTATCATATTTCACGTGC	XhoI
fu-(ΔG287)-919-His	Fwd	CGC<u>GGATCC</u>GGTGGTGGTGGT-CAAAGCAAGAGCATCCAAACC	BamHI
	Rev	CCC<u>AAGCTT</u>-TTCGGGCGGTATTCGGGCTTC	HindIII
fu-(ΔG287)-953-His	Fwd	CGC<u>GGATCC</u>GGTGGTGGTGGT-GCCACCTACAAAGTGGAC	BamHI
	Rev	GCCC<u>AAGCTT</u>-TTGTTTGGCTGCCTCGAT	HindIII
fu-(ΔG287)-961-His	Fwd	CGC<u>GGATCC</u>GGTGGTGGTGGT-ACAAGCGACGACG	BamHI
	Rev	GCCC<u>AAGCTT</u>-CCACTCGTAATTGACGCC	HindIII
fu-(ΔG287)-Orf46.1-His	Fwd	CGC<u>GGATCC</u>GGTGGTGGTGGT-TCAGATTTGGCAAACGATTC	BamHI
	Rev	CCC<u>AAGCTT</u>-CGTATCATATTTCACGTGC	HindIII
fu-(ΔG287-919)-Orf46.1-His	Fwd	CCC<u>AAGCTT</u>GGTGGTGGTGGTGGT-TCAGATTTGGCAAACGATTC	HindIII
	Rev	CCC<u>GCTCGAG</u>-CGTATCATATTTCACGTGC	XhoI
fu-(ΔG287-Orf46.1)-919-His	Fwd	CCC<u>AAGCTT</u>GGTGGTGGTGGTGGT-CAAAGCAAGAGCATCCAAACC	HindIII
	Rev	CCC<u>GCTCGAG</u>-CGGGCGGTATTCGGGCTT	XhoI
fu ΔG287(394.98)-...	Fwd	CGCGGATCC<u>GCTAGC</u>-CCCGATGTTAAATCGGC	NheI
	Rev	CGG<u>GGATCC</u>-ATCCTGCTCTTTTTTGCCGG	BamHI
fu Orf1-(Orf46.1)-His	Fwd	CGCGGATCC<u>GCTAGC</u>-GGACACACTTATTTCGGCATC	NheI
	Rev	CGCGGATCC-CCAGCGGTAGCCTAATTTGAT	
fu (Orf1)-Orf46.1-His	Fwd	CGC<u>GGATCC</u>GGTGGTGGTGGT-TCAGATTTGGCAAACGATTC	BamHI
	Rev	CCC<u>AAGCTT</u>-CGTATCATATTTCACGTGC	HindIII
fu (919)-Orf46.1-His	Fwd1	GCGGC<u>GTCGAC</u>GGTGGCGGAGGCACTGGATCCTCAG	SalI
	Fwd2	GGAGGCACTGGATCCTCAGATTTGGCAAACGATTC	
	Rev	CCC<u>GCTCGAG</u>-CGTATCATATTTCACGTGC	XhoI
Fu (orf46)-287-His	Fwd	CGG<u>GGATCC</u>GGGGGCGGCGGTGGCG	BamHI
	Rev	CCC<u>AAGCTT</u>ATCCTGCTCTTTTTTGCCGGC	HindIII
Fu (orf46)-919-His	Fwd	CGC<u>GGATCC</u>GGTGGTGGTGGTCAAAGCAAGAGCATCCAAACC	BamHI
	Rev	CCC<u>AAGCTT</u>CGGGCGGTATTCGGGCTTC	HindIII

Fu (orf46-919)-287-His	Fwd	CCCCAAGCTTGGGGGCGGCGGTGGCG	HindIII
	Rev	CCCGCTCGAGATCCTGCTCTTTTTTGCCGGC	XhoI
Fu (orf46-287)-919-His	Fwd	CCCAAGCTTGGTGGTGGTGGTGGTCAAAGCAAGAGCATCCAAACC	HindIII
	Rev	CCCGCTCGAGCGGGCGGTATTCGGGCTT	XhoI
(ΔG741)-961c-His	Fwd1 Fwd2	GGAGGCACTGGATCCGCAGCCACAAACGACGACGA GCGGCCTCGAG-GGTGGCGGAGGCACTGGATCCGCAG	XhoI
	Rev	CCCGCTCGAG-ACCCAGCTTGTAAGGTTG	XhoI
(ΔG741)-961-His	Fwd1 Fwd2	GGAGGCACTGGATCCGCAGCCACAAACGACGACGA GCGGCCTCGAG-GGTGGCGGAGGCACTGGATCCGCAG	XhoI
	Rev	CCCGCTCGAG-CCACTCGTAATTGACGCC	XhoI
(ΔG741)-983-His	Fwd	GCGGCCTCGAG-GGATCCGGCGGAGGCGGCACTTCTGCG	XhoI
	Rev	CCCGCTCGAG-GAACCGGTAGCCTACG	XhoI
(ΔG741)-orf46.1-His	Fwd1 Fwd2	GGAGGCACTGGATCCTCAGATTTGGCAAACGATTC GCGGCGTCGACGGTGGCGGAGGCACTGGATCCTCAGA	SalI
	Rev	CCCGCTCGAG-CGTATCATATTTCACGTGC	XhoI
(ΔG983)-741(MC58) -His	Fwd	GCGGCCTCGAG-GGATCCGGAGGGGGTGGTGTCGCC	XhoI
	Rev	CCCGCTCGAG-TTGCTTGGCGGCAAG	XhoI
(ΔG983)-961c-His	Fwd1 Fwd2	GGAGGCACTGGATCCGCAGCCACAAACGACGACGA GCGGCCTCGAG-GGTGGCGGAGGCACTGGATCCGCAG	XhoI
	Rev	CCCGCTCGAG-ACCCAGCTTGTAAGGTTG	XhoI
(ΔG983)-961-His	Fwd1 Fwd2	GGAGGCACTGGATCCGCAGCCACAAACGACGACGA GCGGCCTCGAG-GGTGGCGGAGGCACTGGATCCGCAG	XhoI
	Rev	CCCGCTCGAG-CCACTCGTAATTGACGCC	XhoI
(ΔG983)- Orf46.1-His	Fwd1 Fwd2	GGAGGCACTGGATCCTCAGATTTGGCAAACGATTC GCGGCGTCGACGGTGGCGGAGGCACTGGATCCTCAGA	SalI
	Rev	CCCGCTCGAG-CGTATCATATTTCACGTGC	XhoI

* This primer was used as a Reverse primer for all the C terminal fusions of **287** to the His-tag.

§ Forward primers used in combination with the 287-His Reverse primer.

NB – All PCR reactions use strain 2996 unless otherwise specified (*e.g.* strain MC58)

In all constructs starting with an ATG not followed by a unique *Nhe*I site, the ATG codon is part of the *Nde*I site used for cloning. The constructs made using *Nhe*I as a cloning site at the 5' end (*e.g.* all those containing 287 at the N-terminus) have two additional codons (GCT AGC) fused to the coding sequence of the antigen.

Preparation of chromosomal DNA templates

N.meningitidis strains 2996, MC58, 394.98, 1000 and BZ232 (and others) were grown to exponential phase in 100ml of GC medium, harvested by centrifugation, and resuspended in 5ml buffer (20% w/v sucrose, 50mM Tris-HCl, 50mM EDTA, pH8). After 10 minutes incubation on ice, the bacteria were lysed by adding 10ml of lysis solution (50mM NaCl, 1% Na-Sarkosyl, 50µg/ml Proteinase K), and the suspension incubated at 37°C for 2 hours. Two

phenol extractions (equilibrated to pH 8) and one $CHCl_3$/isoamylalcohol (24:1) extraction were performed. DNA was precipitated by addition of 0.3M sodium acetate and 2 volumes of ethanol, and collected by centrifugation. The pellet was washed once with 70%(v/v) ethanol and redissolved in 4.0ml TE buffer (10mM Tris-HCl, 1mM EDTA, pH 8.0). The DNA concentration was measured by reading OD_{260}.

PCR Amplification

The standard PCR protocol was as follows: 200ng of genomic DNA from 2996, MC581000, or BZ232 strains or 10ng of plasmid DNA preparation of recombinant clones were used as template in the presence of 40μM of each oligonucletide primer, 400-800 μM dNTPs solution, 1x PCR buffer (including 1.5mM $MgCl_2$), 2.5 units *TaqI* DNA polymerase (using Perkin-Elmer AmpliTaQ, Boerhingher Mannheim Expand™ Long Template).

After a preliminary 3 minute incubation of the whole mix at 95°C, each sample underwent a two-step amplification: the first 5 cycles were performed using the hybridisation temperature that excluded the restriction enzyme tail of the primer (T_{m1}). This was followed by 30 cycles according to the hybridisation temperature calculated for the whole length oligos (T_{m2}). Elongation times, performed at 68°C or 72°C, varied according to the length of the Orf to be amplified. In the case of Orf1 the elongation time, starting from 3 minutes, was increased by 15 seconds each cycle. The cycles were completed with a 10 minute extension step at 72°C.

The amplified DNA was either loaded directly on a 1% agarose gel. The DNA fragment corresponding to the band of correct size was purified from the gel using the Qiagen Gel Extraction Kit, following the manufacturer's protocol.

Digestion of PCR fragments and of the cloning vectors

The purified DNA corresponding to the amplified fragment was digested with the appropriate restriction enzymes for cloning into pET-21b+, pET22b+ or pET-24b+. Digested fragments were purified using the QIAquick PCR purification kit (following the manufacturer's instructions) and eluted with either H_2O or 10mM Tris, pH 8.5. Plasmid vectors were digested with the appropriate restriction enzymes, loaded onto a 1.0% agarose gel and the band corresponding to the digested vector purified using the Qiagen QIAquick Gel Extraction Kit.

Cloning

The fragments corresponding to each gene, previously digested and purified, were ligated into pET21b+, pET22b+ or pET-24b+. A molar ratio of 3:1 fragment/vector was used with T4 DNA ligase in the ligation buffer supplied by the manufacturer.

Recombinant plasmid was transformed into competent *E.coli* DH5 or HB101 by incubating the ligase reaction solution and bacteria for 40 minutes on ice, then at 37°C for 3 minutes. This was followed by the addition of 800µl LB broth and incubation at 37°C for 20 minutes. The cells were centrifuged at maximum speed in an Eppendorf microfuge, resuspended in approximately 200µl of the supernatant and plated onto LB ampicillin (100mg/ml) agar.

Screening for recombinant clones was performed by growing randomly selected colonies overnight at 37°C in 4.0ml of LB broth + 100µg/ml ampicillin. Cells were pelleted and plasmid DNA extracted using the Qiagen QIAprep Spin Miniprep Kit, following the manufacturer's instructions. Approximately 1µg of each individual miniprep was digested with the appropriate restriction enzymes and the digest loaded onto a 1-1.5% agarose gel (depending on the expected insert size), in parallel with the molecular weight marker (1kb DNA Ladder, GIBCO). Positive clones were selected on the basis of the size of insert.

Expression

After cloning each gene into the expression vector, recombinant plasmids were transformed into *E.coli* strains suitable for expression of the recombinant protein. 1µl of each construct was used to transform *E.coli* BL21-DE3 as described above. Single recombinant colonies were inoculated into 2ml LB+Amp (100µg/ml), incubated at 37°C overnight, then diluted 1:30 in 20ml of LB+Amp (100µg/ml) in 100ml flasks, to give an OD_{600} between 0.1 and 0.2. The flasks were incubated at 30°C or at 37°C in a gyratory water bath shaker until OD_{600} indicated exponential growth suitable for induction of expression (0.4-0.8 OD). Protein expression was induced by addition of 1.0mM IPTG. After 3 hours incubation at 30°C or 37°C the OD_{600} was measured and expression examined. 1.0ml of each sample was centrifuged in a microfuge, the pellet resuspended in PBS and analysed by SDS-PAGE and Coomassie Blue staining.

Purification of His-tagged proteins

Various forms of 287 were cloned from strains 2996 and MC58. They were constructed with a C-terminus His-tagged fusion and included a mature form (aa 18-427), constructs with

deletions (Δ1, Δ 2, Δ3 and Δ4) and clones composed of either B or C domains. For each clone purified as a His-fusion, a single colony was streaked and grown overnight at 37°C on a LB/Amp (100 μg/ml) agar plate. An isolated colony from this plate was inoculated into 20ml of LB/Amp (100 μg/ml) liquid medium and grown overnight at 37°C with shaking. The overnight culture was diluted 1:30 into 1.0 L LB/Amp (100 μg/ml) liquid medium and allowed to grow at the optimal temperature (30 or 37°C) until the OD_{550} reached 0.6-0.8. Expression of recombinant protein was induced by addition of IPTG (final concentration 1.0mM) and the culture incubated for a further 3 hours. Bacteria were harvested by centrifugation at 8000*g* for 15 min at 4°C. The bacterial pellet was resuspended in 7.5 ml of either (i) cold buffer A (300 mM NaCl, 50 mM phosphate buffer, 10 mM imidazole, pH 8.0) for soluble proteins or (ii) buffer B (10mM Tris-HCl, 100 mM phosphate buffer, pH 8.8 and, optionally, 8M urea) for insoluble proteins. Proteins purified in a soluble form included 287-His, Δ1, Δ2, Δ3 and Δ4287-His, Δ4287MC58-His, 287c-His and 287cMC58-His. Protein 287bMC58-His was insoluble and purified accordingly. Cells were disrupted by sonication on ice four times for 30 sec at 40W using a Branson sonifier 450 and centrifuged at 13000xg for 30 min at 4°C. For insoluble proteins, pellets were resuspended in 2.0 ml buffer C (6 M guanidine hydrochloride, 100 mM phosphate buffer, 10 mM Tris- HCl, pH 7.5 and treated with 10 passes of a Dounce homogenizer. The homogenate was centrifuged at 13000*g* for 30 min and the supernatant retained. Supernatants for both soluble and insoluble preparations were mixed with 150μl Ni^{2+}-resin (previously equilibrated with either buffer A or buffer B, as appropriate) and incubated at room temperature with gentle agitation for 30 min. The resin was Chelating Sepharose Fast Flow (Pharmacia), prepared according to the manufacturer's protocol. The batch-wise preparation was centrifuged at 700*g* for 5 min at 4°C and the supernatant discarded. The resin was washed twice (batch-wise) with 10ml buffer A or B for 10 min, resuspended in 1.0 ml buffer A or B and loaded onto a disposable column. The resin continued to be washed with either (i) buffer A at 4°C or (ii) buffer B at room temperature, until the OD_{280} of the flow-through reached 0.02-0.01. The resin was further washed with either (i) cold buffer C (300mM NaCl, 50mM phosphate buffer, 20mM imidazole, pH 8.0) or (ii) buffer D (10mM Tris-HCl, 100mM phosphate buffer, pH 6.3 and, optionally, 8M urea) until OD_{280} of the flow-through reached 0.02-0.01. The His-fusion protein was eluted by addition of 700μl of either (i) cold elution buffer A (300 mM NaCl, 50mM phosphate buffer, 250 mM imidazole, pH 8.0) or (ii) elution buffer B (10 mM Tris-HCl, 100 mM phosphate buffer, pH 4.5 and, optionally, 8M urea) and fractions

collected until the OD_{280} indicated all the recombinant protein was obtained. 20μl aliquots of each elution fraction were analysed by SDS-PAGE. Protein concentrations were estimated using the Bradford assay.

Renaturation of denatured His-fusion proteins.

Denaturation was required to solubilize 287bMC8, so a renaturation step was employed prior to immunisation. Glycerol was added to the denatured fractions obtained above to give a final concentration of 10% v/v. The proteins were diluted to 200 μg/ml using dialysis buffer I (10% v/v glycerol, 0.5M arginine, 50 mM phosphate buffer, 5.0 mM reduced glutathione, 0.5 mM oxidised glutathione, 2.0M urea, pH 8.8) and dialysed against the same buffer for 12-14 hours at 4°C. Further dialysis was performed with buffer II (10% v/v glycerol, 0.5M arginine, 50mM phosphate buffer, 5.0mM reduced glutathione, 0.5mM oxidised glutathione, pH 8.8) for 12-14 hours at 4°C. Protein concentration was estimated using the formula:

$$Protein\ (mg/ml) = (1.55 \times OD_{280}) - (0.76 \times OD_{260})$$

Immunization

Balb/C mice were immunized with antigens on days 0, 21 and 35 and sera analyzed at day 49.

Sera analysis – ELISA

The acapsulated MenB M7 and the capsulated strains were plated on chocolate agar plates and incubated overnight at 37°C with 5% CO_2. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into Mueller-Hinton Broth (Difco) containing 0.25% glucose. Bacterial growth was monitored every 30 minutes by following OD_{620}. The bacteria were let to grow until the OD reached the value of 0.4-0.5. The culture was centrifuged for 10 minutes at 4000rpm. The supernatant was discarded and bacteria were washed twice with PBS, resuspended in PBS containing 0.025% formaldehyde, and incubated for 1 hour at 37°C and then overnight at 4°C with stirring. 100μl bacterial cells were added to each well of a 96 well Greiner plate and incubated overnight at 4°C. The wells were then washed three times with PBT washing buffer (0.1% Tween-20 in PBS). 200μl of saturation buffer (2.7% polyvinylpyrrolidone 10 in water) was added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. 200μl of diluted sera (Dilution buffer: 1% BSA, 0.1% Tween-20, 0.1% NaN_3 in PBS) were added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. 100μl of HRP-conjugated rabbit anti-mouse (Dako) serum diluted 1:2000 in dilution buffer were added to each well and the plates were incubated for 90 minutes at 37°C. Wells

were washed three times with PBT buffer. 100μl of substrate buffer for HRP (25ml of citrate buffer pH5, 10mg of O-phenildiamine and 10μl of H_2O_2) were added to each well and the plates were left at room temperature for 20 minutes. 100μl 12.5% H_2SO_4 was added to each well and OD_{490} was followed. The ELISA titers were calculated abitrarely as the dilution of sera which gave an OD_{490} value of 0.4 above the level of preimmune sera. The ELISA was considered positive when the dilution of sera with OD_{490} of 0.4 was higher than 1:400.

Sera analysis – FACS Scan bacteria binding assay

The acapsulated MenB M7 strain was plated on chocolate agar plates and incubated overnight at 37°C with 5% CO_2. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into 4 tubes containing 8ml each Mueller-Hinton Broth (Difco) containing 0.25% glucose. Bacterial growth was monitored every 30 minutes by following OD_{620}. The bacteria were let to grow until the OD reached the value of 0.35-0.5. The culture was centrifuged for 10 minutes at 4000rpm. The supernatant was discarded and the pellet was resuspended in blocking buffer (1% BSA in PBS, 0.4% NaN_3) and centrifuged for 5 minutes at 4000rpm. Cells were resuspended in blocking buffer to reach OD_{620} of 0.05. 100μl bacterial cells were added to each well of a Costar 96 well plate. 100μl of diluted (1:100, 1:200, 1:400) sera (in blocking buffer) were added to each well and plates incubated for 2 hours at 4°C. Cells were centrifuged for 5 minutes at 4000rpm, the supernatant aspirated and cells washed by addition of 200μl/well of blocking buffer in each well. 100μl of R-Phicoerytrin conjugated $F(ab)_2$ goat anti-mouse, diluted 1:100, was added to each well and plates incubated for 1 hour at 4°C. Cells were spun down by centrifugation at 4000rpm for 5 minutes and washed by addition of 200μl/well of blocking buffer. The supernatant was aspirated and cells resuspended in 200μl/well of PBS, 0.25% formaldehyde. Samples were transferred to FACScan tubes and read. The condition for FACScan (Laser Power 15mW) setting were: FL2 on; FSC-H threshold:92; FSC PMT Voltage: E 01; SSC PMT: 474; Amp. Gains 6.1; FL-2 PMT: 586; compensation values: 0.

Sera analysis – bactericidal assay

N. meningitidis strain 2996 was grown overnight at 37°C on chocolate agar plates (starting from a frozen stock) with 5% CO_2. Colonies were collected and used to inoculate 7ml Mueller-Hinton broth, containing 0.25% glucose to reach an OD_{620} of 0.05-0.08. The culture was incubated for approximately 1.5 hours at 37 degrees with shacking until the OD_{620} reached the value of 0.23-0.24. Bacteria were diluted in 50mM Phosphate buffer pH 7.2 containing 10mM $MgCl_2$, 10mM $CaCl_2$ and 0.5% (w/v) BSA (assay buffer) at the working dilution of 10^5 CFU/ml. The total volume of the final reaction mixture was 50 μl with 25 μl

of serial two fold dilution of test serum, 12.5 µl of bacteria at the working dilution, 12.5 µl of baby rabbit complement (final concentration 25%).

Controls included bacteria incubated with complement serum, immune sera incubated with bacteria and with complement inactivated by heating at 56°C for 30'. Immediately after the addition of the baby rabbit complement, 10µl of the controls were plated on Mueller-Hinton agar plates using the tilt method (time 0). The 96-wells plate was incubated for 1 hour at 37°C with rotation. 7µl of each sample were plated on Mueller-Hinton agar plates as spots, whereas 10µl of the controls were plated on Mueller-Hinton agar plates using the tilt method (time 1). Agar plates were incubated for 18 hours at 37 degrees and the colonies corresponding to time 0 and time 1 were counted.

Sera analysis – western blots

Purified proteins (500ng/lane), outer membrane vesicles (5µg) and total cell extracts (25µg) derived from MenB strain 2996 were loaded onto a 12% SDS-polyacrylamide gel and transferred to a nitrocellulose membrane. The transfer was performed for 2 hours at 150mA at 4°C, using transfer buffer (0.3% Tris base, 1.44% glycine, 20% (v/v) methanol). The membrane was saturated by overnight incubation at 4°C in saturation buffer (10% skimmed milk, 0.1% Triton X100 in PBS). The membrane was washed twice with washing buffer (3% skimmed milk, 0.1% Triton X100 in PBS) and incubated for 2 hours at 37°C with mice sera diluted 1:200 in washing buffer. The membrane was washed twice and incubated for 90 minutes with a 1:2000 dilution of horseradish peroxidase labelled anti-mouse Ig. The membrane was washed twice with 0.1% Triton X100 in PBS and developed with the Opti-4CN Substrate Kit (Bio-Rad). The reaction was stopped by adding water.

The OMVs were prepared as follows: *N. meningitidis* strain 2996 was grown overnight at 37 degrees with 5% CO_2 on 5 GC plates, harvested with a loop and resuspended in 10 ml of 20mM Tris-HCl pH 7.5, 2 mM EDTA. Heat inactivation was performed at 56°C for 45 minutes and the bacteria disrupted by sonication for 5 minutes on ice (50% duty cycle, 50% output , Branson sonifier 3 mm microtip). Unbroken cells were removed by centrifugation at 5000*g* for 10 minutes, the supernatant containing the total cell envelope fraction recovered and further centrifuged overnight at 50000*g* at the temperature of 4°C . The pellet containing the membranes was resuspended in 2% sarkosyl, 20mM Tris-HCl pH 7.5, 2 mM EDTA and incubated at room temperature for 20 minutes to solubilise the inner membranes. The suspension was centrifuged at 10000*g* for 10 minutes to remove aggregates, the supernatant was further centrifuged at 50000*g* for 3 hours. The pellet, containing the outer membranes

was washed in PBS and resuspended in the same buffer. Protein concentration was measured by the D.C. Bio-Rad Protein assay (Modified Lowry method), using BSA as a standard.

Total cell extracts were prepared as follows: *N. meningitidis* strain 2996 was grown overnight on a GC plate, harvested with a loop and resuspended in 1ml of 20mM Tris-HCl. Heat inactivation was performed at 56°C for 30 minutes.

CLAIMS

1. A method for the simultaneous heterologous expression of two or more proteins of the invention, in which (a) two or more proteins of the invention are fused.

2. The method of claim 23, in which the two or more proteins are: (a) 919 and 287; (b) 953 and 287; (c) 287 and ORF46.1; (d) ORF1 and ORF46.1; (e) 919 and ORF46.1; (f) ORF46.1, 287 and 919; (g) 919 and 519; and (h) ORF97 and 225.

3. The method of claim 24, in which 287 is at the C-terminal end of protein (a), (b) or (c).

4. The method of any preceding claim, in which the expression is in an *E.coli* host.

5. A protein expressed by the method of any preceding claim.

6. A hybrid protein of formula NH_2-A—B-COOH, wherein A and B are different Neisserial proteins.

7. The protein of claim 6, wherein A and B are each selected from orf1, orf4, orf25, orf40, orf46, orf83, 233, 287, 2921, 564, 687, 741, 907, 919, 953, 961 and 983.

8. The protein of claim 7, wherein A and B are each selected from ORF46.1, 287, 741, 919, 953, 961 and 983.

9. The protein of claim 8, wherein at least one of said ORF46.1, 287, 741, 919, 953, 961 and 983 is used in essentially full-length form

10. The protein of claim 8 or claim 9, wherein at least one of said ORF46.1, 287, 741, 919, 953, 961 and 983 has a deletion.

11. The protein of claim 10, wherein A and/or B has a poly-glycine deletion ('ΔG').

12. The protein of claim 11, wherein A and/or B is ΔG-287, ΔGTbp2, ΔG741, or ΔG983.

13. The protein of claim 10, wherein A and/or B is a truncated protein.

14. The protein of claim 13, wherein A and/or B is Δ1-287, Δ2-287, Δ3-287 or Δ4-287.

15. The protein of claim 10, wherein a domain of A and/or B is deleted.

16. The protein of claim 15, wherein A and/or B is 287B, 287C, 287BC, ORF461-433, ORF46433-608, ORF46, or 961c.

17. The protein of claim 6, wherein A and B are: (a) 919 and 287; (b) 953 and 287; (c) 287 and ORF46.1; (d) ORF1 and ORF46.1; (e) 919 and ORF46.1; (f) ORF46.1 and 919; (g) 919 and 519; or (h) ORF97 and 225.

18. The protein of claim 17, wherein the protein is ΔG287-919, ΔG287-953, ΔG287-961, ΔG983-ORF46.1, ΔG983-741, ΔG983-961, ΔG983-961C, ΔG741-961, ΔG741-961C, ΔG741-983, ΔG741-ORF46.1, ORF46.1-741, ORF46.1-961, ORF46.1-961C, 961-ORF46.1, 961-741, 961-983, 961C-ORF46.1, 961C-741, 961C-983, 961CL–ORF46.1, 961CL-741, or 961CL-983.

19. The protein of claim 8, wherein A or B is 287.

20. The protein of claim 19, wherein B is 287

21. The protein of claim 19, wherein A is ΔG-287

22. The protein of claim 21, wherein B is ORF46, 919, 953 or 961.

23. The protein of any one of claims 19 to 22, wherein 287 is from strain 2996 or 394/98.

24. The protein of claim 8, wherein A is 961.

25. The protein of any one of claims 6 to 24, wherein A and B are from the same strain.

26. The protein of any one of claims 6 to 24, wherein A and B are joined directly

27. The protein of any one of claims 6 to 24, wherein A and B are joined via a linker peptide.

28. The protein of claim 27, wherein the linker peptide is a poly-glycine linker, with the proviso that B is not a ΔG protein.

FIGURE 1 — ΔG287—919

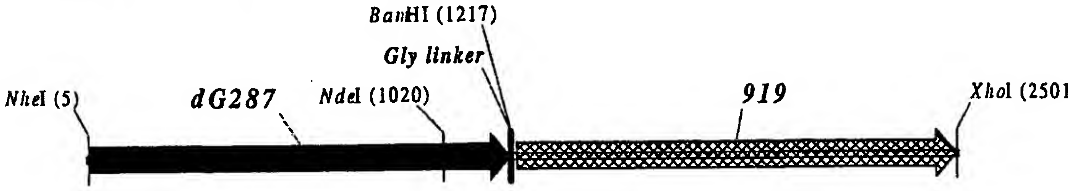


FIGURE 2 — ΔG287—953

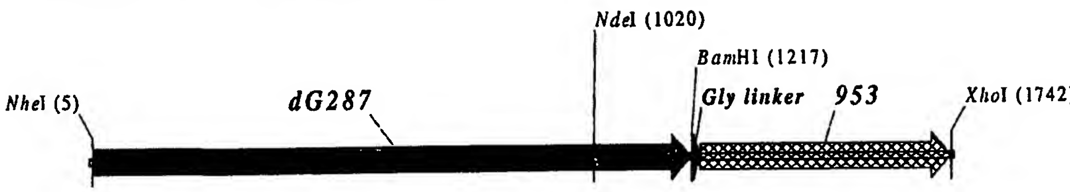


FIGURE 3 — ΔG287—961

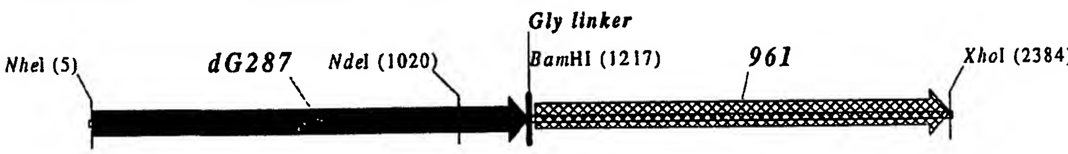


FIGURE 4 — ΔG287NZ—919

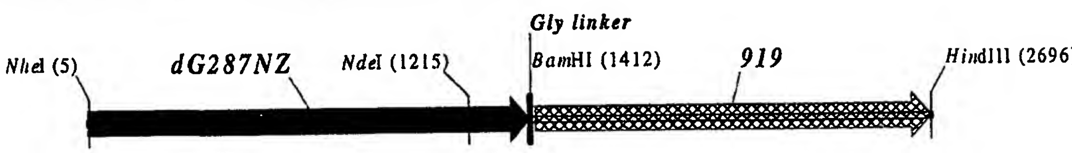


FIGURE 5 — ΔG287NZ—953

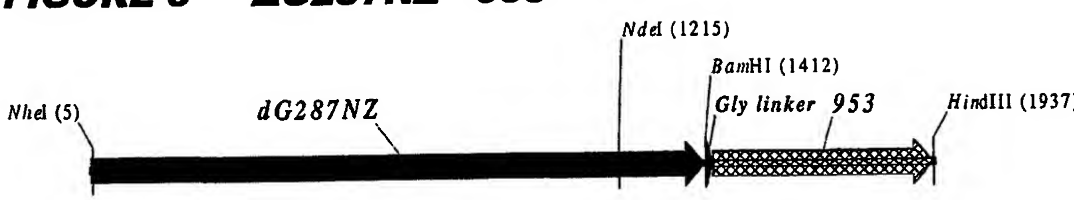


FIGURE 6 — ΔG287NZ—961

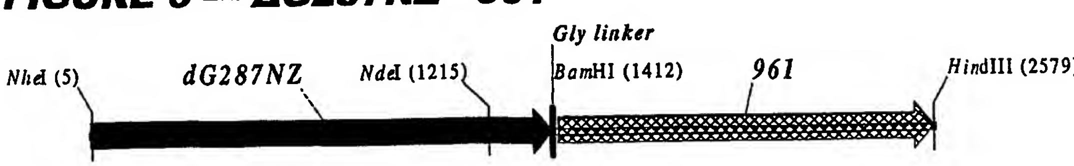


FIGURE 7 — ΔG983-ORF46.1

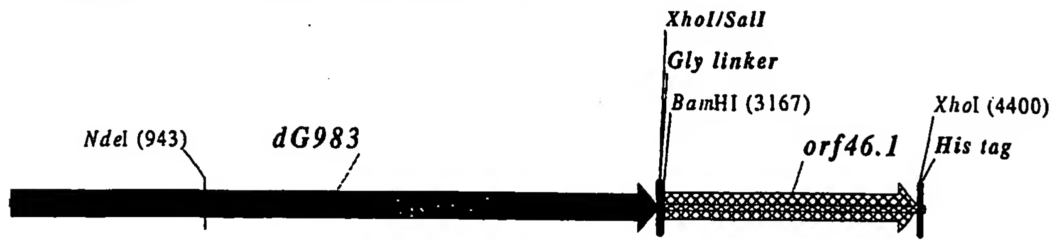


FIGURE 8 — ΔG983-741

FIGURE 9 — ΔG983-961

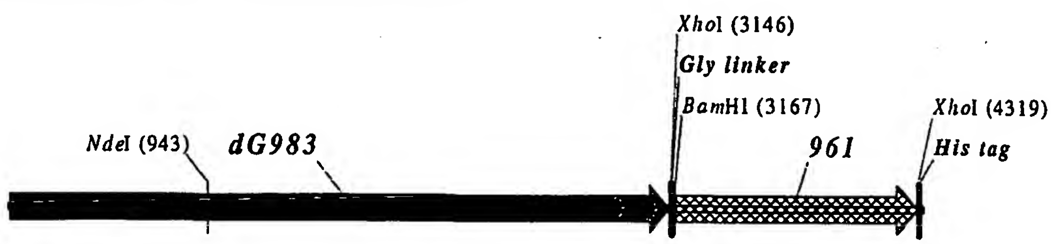


FIGURE 10 — ΔG983-961c

FIGURE 11 — ΔG741-961

FIGURE 12 — ΔG741-961c

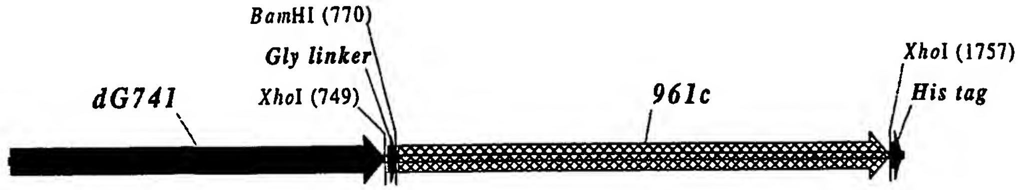


FIGURE 13 — ΔG741-983

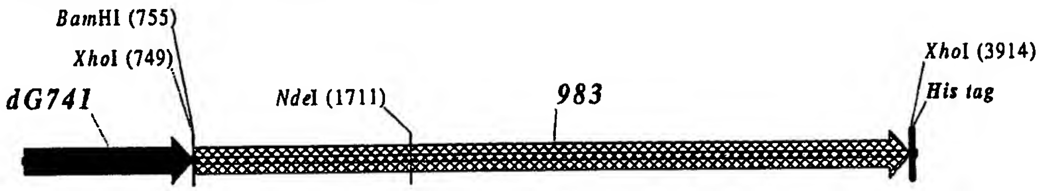


FIGURE 14 — ΔG741-ORF46.1

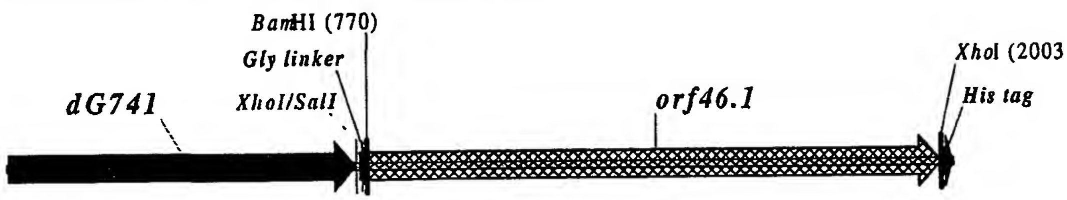


FIGURE 15 — ORF46.1-741

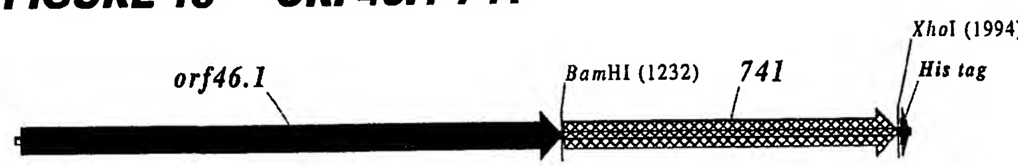


FIGURE 16 — ORF46.1-961

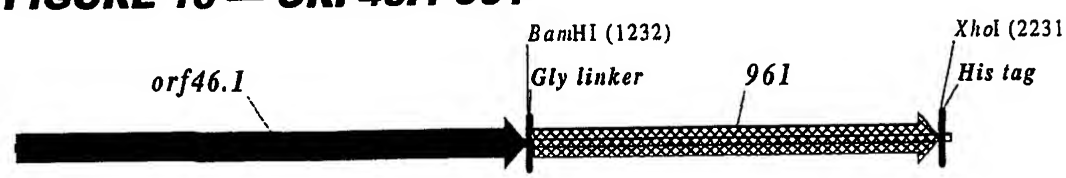


FIGURE 17 — ORF46.1—961c

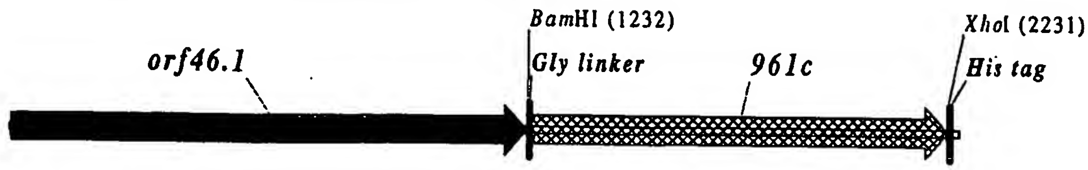


FIGURE 18 — 961-ORF46.1

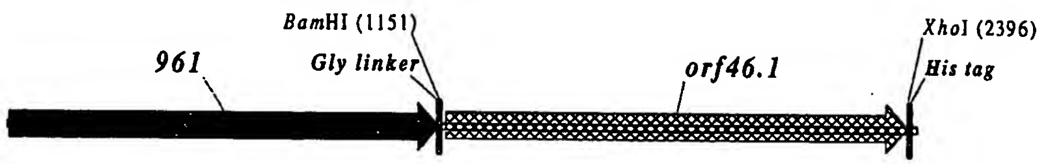


FIGURE 19 — 961-741

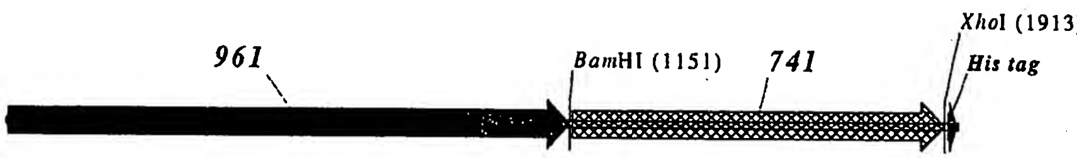


FIGURE 20 — 961-983

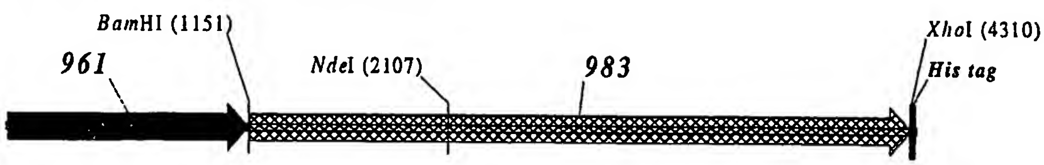


FIGURE 21 — 961c-ORF46.1

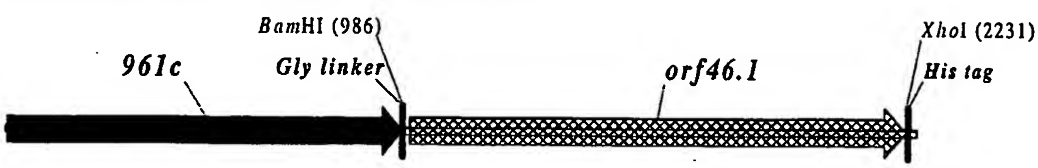
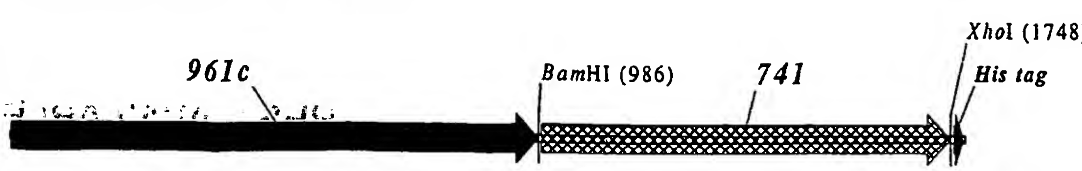


FIGURE 22 — 961c-741



BEST AVAILABLE COPY

FIGURE 23 — 961c-983

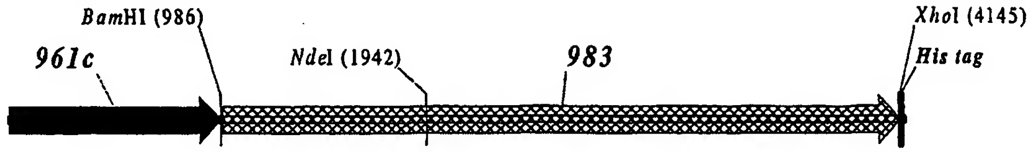


FIGURE 24 — 961cL-ORF46.1

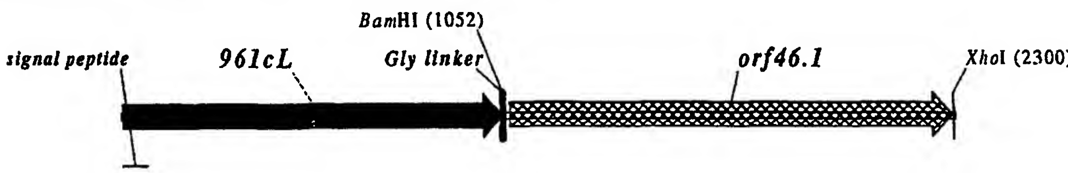


FIGURE 25 — 961cL-741

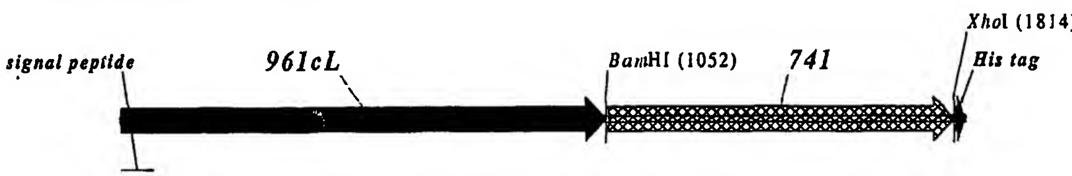


FIGURE 26 — 961cL-983